DOCUMENT RESUME

ED 109 139	TM 004 516
TITLE	Summary of Research Projects 1973-1974 [at Educational Testing Service].
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-R-54
PUB DATE	Sep 74
NOTE	202p.; For related documents, see ED 056 074 and 076
EDRS PRICE	MF-$0.76 HC-$10.78 PLUS POSTAGE
DESCRIPTORS	*Annotated Bibliographies; Career Education; Cognitive Processes; Computer Science; Cultural Factors; Early Childhood Education; *Educational Research; Elementary Secondary Education; Graduate Study; Higher Education; Human Development; Learning; Measurement Techniques; Personality; Professional Education; Program Evaluation; Research Methodology; *Research Reviews (Publications); Social Behavior; Student Evaluation
IDENTIFIERS	*Educational Testing Service

ABSTRACT

Research projects that have been carried out at Educational Testing Service during the period July 1, 1973 through June 30, 1974 are summarized. The summaries are classified as research on: human development, learning and cognition, personality, social behavior, early and primary education, secondary education, higher education, graduate and professional education, cultural diversity and the disadvantaged, career and vocation, evaluation of educational systems, psychometric theory, methodological research, and computer studies. Summaries of projects carried out in previous years can be found in related documents ED 056 074 and 076. (DEP)

****

* Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original. *

****

ED109139

SUMMARY OF RESEARCH PROJECTS

1973 - 1974

(Report #54)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
DOROTHY URBAN
ETS
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

TM 004 516

Educational Testing Service
Princeton, New Jersey
September 1974

Copyright © 1974. Educational Testing Service. All rights reserved.

FOREWORD

The "Summary of Research Projects" reports on research carried on at ETS during the period July 1, 1973, through June 30, 1974.

Questions concerning the studies reported herein should be directed to the authors. If a report on a study is available, this will be indicated by a number appearing in the body of the write-up. Please use this number when requesting a report.

TABLE OF CONTENTS

I. PSYCHOLOGICAL RESEARCH

A. Human Development

Baker, A. H.

Developmental Changes in Self/World Differentiation 149-151

Developmental Model of Cognitive Style 126-129

Brooks, J.

The Effect of Time on Attachment as Measured in a Free Play Situation (with M. Lewis) 157

Infant's Responses to Pictures of Mother, Self, and Other (with M. Lewis) 173

Social Perception in Infancy (with M. Lewis). 155

Visual Self Recognition in Infancy (with M. Lewis). 174

Emmerich, W.

Differentiation and Development of Social Norms 137-138

Lewis, M.

Assessment of Infant Socioemotional Development (with L. Michalson) 169

The Determinants of Two-Year-Old Children's Responses to Brief Separations from the Mother (with M. Weinraub) . . . 43

Developmental Study of Attention within the First Two Years of Life (with E. Scott) 122-124

Evaluation of Infant Intelligence: Infant Intelligence Scores--True or False? (with H. McGurk) 125

The Integration of Faces and Voices in Infancy (with L. Rosenwein and H. McGurk) 130

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development (with C. Wilson) 131-133

Lewis, M. (continued)

Mothers and Fathers, Girls and Boys: Attachment Behavior in the First Two Years of Life (with M. Weinraub and P. Ban) 134-135

Mother-Infant Interaction (with S. Lee-Painter) 158-159

Mother-Infant Interaction from the First Day to the Third Month of Life (with S. Lee-Painter) 136

Peer Group Interaction in Infancy (with J. Brooks, L. Michalson, and G. Young) 154

B. Learning and Cognition

Carroll, J.

International Study of Educational Attainment in French as a Foreign Language. 152-153

Studies in the Psychology of Language 80-82

Study of Aptitude-Achievement Relationships 77

Freedle, R.

Psycholinguistic Theory and Models. 83-84

Goodenough, D.

Short-Term Memory Storage during Sleep (with P. Oltman) 48

Hale, G.

Developmental Studies of Selective Attention. 170-171

Harman, H. H.

Reference Measures for Cognitive and Personality Factors 61

Ward, W.

Creativity, Learning, and Attention 85

Development of Cognition and Memory 86

Individual Difference Parameters in Recognition Memory. 87

C. Personality

Damarin, F.

Leveling-Sharpening as a Cognitive Style (with V. McKenna and S. Messick) 67

Frederiksen, N.

Development of Provisional Criteria for the Study of Scientific Creativity (with W. Ward) 26

Effects of Models on Creative Performance in Relation to Anxiety, Ability, and Sex (with F. Evans) 148

Goodenough, D.

The Effect of Stress on Dream Affect (with H. Witkin) 44

Genetic Factors in the Development of Field Dependence and Spatial Abilities: Intercorrelations Among Family Members (with H. Witkin) 45

Genetic Factors in the Development of Field Dependence and Spatial Abilities: X-Linkage Study (with H. Witkin and I. Olkin) 46

Respiratory Correlates of Emotional States (with P. Oltman) 47

Jackson, D.

The Effects of Detoxification on Psychopathology for Alcoholics as Measured by the Differential Personality Inventory 88

Kogan, N.

Longitudinal Study of Creativity 141

Messick, S.

Cognition of Affect 142

Cognitive Styles and Attitudinal Consistency 146

Cognitive Styles and Psychopathology 140

Dimensions of Perceived Item Similarity 66

Perceptual Factors in Cognition and Personality (with J. French) 62

Messick, S. (continued)

- Personality Organization in Cognitive and Social Processes 68-74

Oltman, P.

- Analytical Ability in Verbal Functioning (with H. Witkin). 49
- Cognitive Style as a Factor in Dyadic Interaction Under Conditions of Initial Conflict (with D. Goodenough and H. Witkin). 50
- Laterality Study. 51
- Psychological Differentiation as Related to Antisocial Behavior, Impulsivity, and Endocrine Functioning in a Prison Population (with H. Witkin and D. Owen). 57
- Studies of Young Adults with Uneven Cognitive Development (with H. Witkin). 52

Owen, D.

- Relationship of Height to Criteria of Antisocial Behavior . . . 58

Stricker, L.

- Retest Changes on Personality Scales. 75

Witkin, H.

- Cognitive Patterning in Sighted Children with Bilateral Retinoblastoma (with F. Friedman) 53
- Cross-cultural Studies of the Development of Psychological Differentiation 54
- Genetic Factors in Field Dependence and in Aggressive Behavior: Studies of Sex Chromosome Anomalies (with D. Goodenough and D. Owen). 59
- Modes of Representation of Presleep Experiences in Dreams (with D. Goodenough). 55
- Sex Chromosome Anomalies in a Male Prison Population (with P. Oltman and D. Owen). 60
- Teacher-Student Interaction, Student Achievement and Mutual Attraction between Teacher and Students as a Function of Match or Mismatch in Cognitive Style (with D. Goodenough, P. Oltman, C. A. Moore, and F. McDonald). 56

D. Social Behavior

Kogan, N.

Determinants of Risk Taking for Others (with A. Teger). 144

Dimensions and Categories of Physiognomic Judgment. 143

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (with M. Hermann) . . . 76

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts 145

II. EDUCATIONAL APPLICATIONS

A. Early and Primary Education

Anderson, S.

CIRCUS. 10

Chittenden, E.

Study of Teacher Change in Open Educational Settings. 179

Emmerich, W.

Preschool Teachers' Beliefs on Effective Teaching Techniques and Their Relationships to Pupil Characteristics 29

Norton, D.

Evaluation of Preschool for Urban Children Project. 8

Minneapolis Pyramids Sibling Study, Phase III 9

Powers, D.

Evaluation of the New Approach Method 36

Sigel, I.

Center for Child Care Research (with R. Cocking and B. Bryant). 139

Trismen, D.

A Descriptive and Analytic Study of Compensatory Reading Programs. 38

B. Secondary Education

Diederich, P.

Analysis of Data on the Cooperative Literature Tests. 78

Book on Measurement of English Language Arts. 79

Evans, F.

SAT Speededness Study 6

Schrader, W. B.

Developing 1974 Norms for the Preliminary Scholastic Aptitude Test (with R. Jackson) 7

Effect of Moderate Departures from Parallelism on the Precision of SAT Equating 4

Stern, J.

College Board Activity and Participation Statistics 1

C. Higher Education

Centra, J.

Departmental Characteristics and Student Academic Growth. . . . 177

The Effectiveness of Student Feedback in Modifying College Instruction 182

Self-Ratings of College Teachers: A Comparison to Student Ratings 183

Student Points of View in Ratings of College Instruction (with R. Linn). 180

Evans, F.

GRE-Quantitative Coaching Study 23-24

Warren, J.

Study of Bachelor's Degrees 156

D. Graduate and Professional Education

Campbell, J.

Evaluation of Program for Training Dentists in the Out-of-Hospital Care of the Handicapped (with R. Flaugher, B. Esser, and J. Knapp). 41

A Survey of Skills and Abilities Needed for Graduate Study in Business (with P. Casserly). 42

Carlson, A.

Development of Graduate Student Rating Scales (with R. Reilly). 18

Pitcher, B.

Student Subgroups Validity Study. 17

Validity Study of LSAT Part-Score Reporting 16

Reilly, R.

Graduate Performance by Critical Incident 20

A Study of Empirical Option Weighting in the LSAT 13

Schrader, W.

Extension of the Study of the Possibility of Curvilinearity in Law School Validity Study Data (with B. Pitcher) 14

Witkin, H.

The Role of Cognitive Style in Students' Academic Evolution from High School Through Graduate and Professional School (with P. Oltman, F. Friedman, and D. Goodenough). 25

E. Cultural Diversity and the Disadvantaged

Campbell, J.

Test and Job Performance of Various Subgroups within Specified Occupations (with L. Crooks, D. Rock, and M. Mahoney) 181

Campbell, J. (continued)

Word Associations of Black and White College Students 21

Emmerich, W.

Development of Sex-Role Constancy in Disadvantaged Children. 30

Personality Structure and Development in Headstart Children . . 31

Feldmesser, R.

Survey of Disparities in Educational Achievements 39

Murphy, R.

Survey of Adult Reading Performance 37

Schrader, W. B.

Prediction of Law School Grades for Mexican American and Black American Students (with B. Pitcher) 15

Shipman, V.

ETS-Head Start Longitudinal Study of Disadvantaged Children and Their First School Experiences. 32

Stricker, L.

Social Stratification 147

Warren, J.

Chicano Validity Study. 5

F. Career and Vocation

Campbell, J.

Development of Interest Measurement Instruments for Assignment of Army Officers to Appropriate Positions (with R. Boldt, R. Flaugher, and D. Rock) 172

Centra, J.

Colleagues as Raters of Classroom Instruction 28

Women Doctoral Recipients 22

Echternacht, G.

Development and Validity of an Air Force Vocational Inventory Study (with R. Reilly). 178

Fremer, J.

Handbook for Conducting Task Analyses and Developing Criterion-Referenced Tests of Language Skills 12

Katz, M.

System of Interactive Guidance and Information. 175

Validating Effects of SIGI on Career Decision-Making Processes of Students 176

Schrader, W. B.

Norming and Validity Studies of North Carolina Nursing Equivalency Examinations (with P. E. Vitella) 11

G. Evaluation of Educational Systems

Fremer, J.

National Assessment of Educational Progress 27

Murphy, R.

Accountability Design for the Cinnaminson Township (New Jersey) School District (with D. Powers) 35

Powers, D.

Evaluation of the Virginia Beach Extended School Year Program 34

III. THEORIES, METHODS, AND SYSTEMS RESEARCH

A. Psychometric Theory

Gleser, L.

Classification under the Model of Compound Symmetry 104

Large-Sample Expansions for Null Distributions of Likelihood Criteria (with I. Olkin) 105

Gleser, L. (continued)

Multivariate Statistical Inference under Marginal Structure. I (with I. Olkin) 106

Multivariate Statistical Inference under Marginal Structure. II (with I. Olkin). 107

Testing If Two Measuring Procedures Measure the Same Dimension 108

Gulliksen, H.

Looking Back and Looking Ahead in Psychometrics 94

Kearns, J.

Posterior True Score Distributions. 119

Kristof, W.

A Discussion of Certain Reliability Formulas. 96

Estimation of Reliability and True Score Variance from a Split of a Test into Three Arbitrary Parts. 97

On Accuracy in Reliability Estimation 98

On the Statistics of Reliability Estimation Using Parallel Test Forms 99

Lord, F.

The 'Ability' Scale in Item Characteristic Curve Theory 167

Cheating Tests. 184

Checking Procedures for Estimating Item and Person (Ability) Parameters by Analyzing Artificial Data 160

Design and Evaluation of Novel Tests for Practical Use (with M. Wingersky) 161

Disattenuated Partial Correlation 63

Estimating Item Parameters (with M. Wingersky). 64

Formula Scoring and Number-Right Scoring. 162

Lord, F. (continued)

Individualized Testing and Item Characteristic Curve Theory. 163

Quick Estimates of the Relative Efficiency of Two Tests as a Function of Ability Level. 164

Relative Efficiency of Number-Right and Formula Scores. 168

The Relative Efficiency of Two Tests as a Function of Ability Level. 165

Testing if Two Measuring Procedures Measure the Same Dimension. 65

Variance Stabilizing Transformation of the Stepped-Up Reliability Coefficient 166

Rubin, D.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors (with T. Stroud). 110

Stroud, T.

Comparing Regressions When Measurement Error Variances Are Known 112

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors. 113

Estimating Variance Components by Combining Unbiased Estimates from Independent Data Sets. 114

B. Methodological Research

Boldt, R.

GRE Bayesian Validity Study 19

Donlon, T.

Differential Weighting of Item Distracters in Data Sufficiency Items 3

Gulliksen, H.

Applications of Psychological Scaling Methods 89

Gulliksen, H. (continued)

Characteristic Roots and Vectors Indicating Agreement of Data with Different Scaling Laws 90

Comments on "Educational Values for Measurement Technology: Some Theory and Data". 91

Learning. 92-93

Jamison, D.

Optimal Choice Theory 40

Kearns, J.

Constrained Latent Class Solutions (with D. Thayer) 116

Empirical Bayes Point Estimates of True Score 117

On the Evaluation of Empirical Bayes Point Estimates. 118

A Relative Interaction Criterion for Factor Rotation. 120

Some Bivariate Convergence Theorems 121

Levine, M.

The Generalization Function in Probability Learning Experiments with Infinitely Many Stimuli and Responses. . . . 100

Geometric Interpretations of Magnitude Estimation and Other Direct Psychophysical Measurement Procedures. 101

Nonadditive Analogues of the Basic Mathematics of Additive Measurement. 102

Parameterizing Curves Transformable to the Same Shape (with D. Saxe). 103

Lord, F.

Equating. 2

Power Scores Estimated by Item Characteristic Curves. 95

Rubin, D.

Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study (with T. Stroud) 109

Stroud, T.

Approximate Confidence Intervals and Tests for the Gamma Distribution: A Monte Carlo Evaluation 111

Use of INDSCAL in Assessing a Cerebral Deterioration Pattern in Psychiatric Patients 115

C. Computer Studies

Jamison, D.

Computer and Communication Technologies and Efficiency in Education. 33

* * * * * * * * *

List of Project Directors. xvii

List of Project Directors

Anderson, S., 10
Baker, A. H., 126, 149
Balloun, J., 120
Ban, P., 134
Boldt, R., 19, 172
Brooks, J., 154, 155, 157, 173, 174
Bryant, B., 139
Campbell, J., 21, 41, 42, 172, 181
Carlson, A., 18
Carroll, J., 77, 80, 152
Casserly, P., 42
Centra, J., 22, 28, 177, 180, 182, 183
Chittenden, E., 179
Cliff, N., 70
Cocking, R., 139
Crooks, L., 181
Damarin, F., 67, 70, 73
Dermen, D., 71, 72, 73, 74
Diederich, P., 78, 79
Donlon, T., 3
Echternacht, G., 178
Emmerich, W., 29, 30, 31, 137
Esser, B., 41
Evans, F., 6, 23, 148
Feldmesser, R., 39
Flaugher, R., 41, 172
Frederiksen, N., 26, 148
Freedle, R., 83
Fremer, J., 12, 27
French, J., 62
Friedman, F., 25, 53
Gleser, L., 104, 105, 106 107, 108
Goodenough, D., 25, 44, 45, 46, 47, 48, 50, 55, 56, 59
Gulliksen, H., 89, 90, 91, 92, 94
Hale, G., 170
Harman, H., 61
Hermann, M., 76
Jackson, D., 88
Jackson, R., 7
Jamison, D., 33, 40
Katz, M., 175, 176
Kearns, J., 116, 117, 118, 119, 120, 121
Knapp, J., 41
Kogan, N., 68, 76, 141, 143, 144, 145
Kristof, W., 96, 97, 98, 99
Lee-Painter, S., 136, 158
Levine, M., 100, 101, 102, 103
Lewis, M., 43, 122, 125, 130, 131, 134, 136, 154, 155, 157, 158, 169, 173, 174
Linn, R., 180
Lord, F., 2, 63, 64, 65, 95, 160, 161, 162, 163, 164, 165, 166, 167, 168, 184
Mahoney, M., 181
McDonald, F., 56
McGurk, H., 125, 130
McKenna, V., 67, 70
Messick, S., 62, 66, 67, 68, 140, 142, 146
Michalson, L., 154, 169
Moore, C. A., 56
Murphy, R., 35, 37
Norton, D., 8, 9
Olkin, I., 46, 105, 106, 107
Oltman, P., 25, 47, 48, 49, 50, 51, 52, 56, 57, 60
Owen, D., 57, 58, 59, 60
Pitcher, B., 14, 15, 16, 17
Powers, D., 34, 35, 36
Reilly, R., 13, 18, 20, 178
Rock, D., 172, 181
Rosenwein, L., 130
Rubin, D., 109, 110
Saxe, D., 103
Schiffman, H., 71, 72
Schrader, W. B., 4, 7, 11, 14, 15
Scott, E., 122
Shipman, V., 32
Sigel, I., 139
Stern, J., 1
Stricker, L., 75, 147
Stroud, T., 109, 110, 111, 112, 113, 114, 115
Teger, A., 144
Thayer, D., 116
Trismen, D., 38
Vitella, P., 11
Ward, W., 26, 85, 86, 87
Warren, J., 5, 156
Weinraub, M., 43, 134
Wilson, C., 131
Wingersky, M., 64, 161
Witkin, H., 25, 44, 45, 46, 49, 50, 52, 53, 54, 55, 56, 57, 59, 60
Young, G., 154

College Board Activity and Participation Statistics (CEEB, 130-2)

J. Stern

PURPOSE: This study is concerned with the preparation of selected descriptive statistics on College Board candidates to describe patterns of test participation with particular reference to shifts that may affect score interpretation and program planning.

PROGRESS: Candidates and Tests, 1972-73 was completed and distributed.

MAJOR FINDINGS: Candidate and test volume in the admissions testing program continued to follow a downward trend in 1972-73. Four indices of overall testing activity used in the admissions testing program showed the following changes in 1972-73 over 1971-72 volume: (1) a decrease of 4.1% in number of Scholastic Aptitude Tests administered, (2) a decrease of 12.6% in number of Achievement Test series, (3) a decrease of 6.1% in total test volume, and (4) a decrease of 5.3% in number of registrations. The number of Preliminary Scholastic Aptitude Tests decreased 3.8%.

Equating (CEEB, 130-3) F. Lord

PURPOSE: (1) To derive a true-score equating method similar to the Levine method. (2) To study the effects of chain equating on bias and sampling variance.

MAJOR FINDINGS: The true-score procedure was found to have no cumulative error over a long single equating chain, nor does the sampling variance grow as the chain lengthens. Sampling variances for various complicated equating chains were computed by AUTEST.

Differential Weighting of Item Distracters in Data Sufficiency Items (CEEB, 141-5)

T. Donlon

PURPOSE: Differential weighting of wrong-answer responses to multiple-choice items may provide additional information not currently reflected in the conventional scoring methods. Data sufficiency items would seem to offer particular promise in that the meaning of responses is constant for all items regardless of content. A comparison will be made of the efficiency of an a priori weighting system and a system using the judgments of mathematics specialists.

PROGRESS: The experimental tests were prepared and administered to approximately 1800 freshman students at seven colleges. Judged estimates of option merits have been secured, and grade-point averages have been received from the schools. Preliminary analyses of the data revealed that none of a number of empirically or rationally derived composite scores, based upon different linear combinations of the different response category scores, yielded an alternate forms reliability as high as the Total Rights score alone.

A number of additional analyses are being performed. They include determining the composite weights by canonical correlational analysis, principal components analysis, and multiple regression analysis, using the Mathematics Achievement Test as the criterion. Validities are being determined for those linear combinations that have a reliability greater than the Total Rights score reliability. Related research in this area by Hendrickson, Reilly and Jackson, and Davis is being reviewed to adapt the analyses most appropriately.

Effect of Moderate Departures from Parallelism on the Precision of SAT Equating

W. B. Schrader

(CEEB, 142-12)

PURPOSE: To ascertain the extent to which developed ability tests of verbal and mathematical ability which placed a moderate emphasis on subject-matter proficiency would deviate from strict parallelism to the Scholastic Aptitude Test, in which an effort is made to minimize the importance of specific course preparation. In this study, variation of equating lines for different student groups (e.g., boys and girls) was used as the indicator of deviation from parallelism.

PROGRESS: Experimental tests of verbal and mathematical ability were administered as part of the Scholastic Aptitude Test. Approximately one-half of the students took the verbal experimental items and the other half took the mathematical items. Information supplied by candidates who took various achievement tests, information obtained from the Student Descriptive Questionnaire, and Achievement Test Scores were used in the analysis. Equatings of the experimental test to SAT-Verbal were performed for 16 subgroups and of the experimental test to SAT-Mathematical were performed for 13 subgroups. Correlations of the experimental verbal test with the College Board English Composition and American History tests were obtained as were correlations of the experimental mathematical tests with the College Board Level I and Level II Mathematics tests. A report has been written and issued as PR-73-41.

MAJOR FINDINGS: For the verbal tests, differences in equating lines were considered to be too large to warrant the construction of a general-purpose equivalency test for the experimental test and SAT-Verbal. For the mathematical tests, the differences were somewhat smaller, but separate conversion lines for students with extensive preparation in mathematics and for students with average or less preparation provided a better fit to the data. On the whole, the differences between the experimental tests and SAT were found to be too large to justify the interchangeable use of equated scores.

Chicano Validity Study (CEEB, 142-13) J. Warren

PURPOSE: Mexican-Americans frequently contend that their bilingual and bicultural background prevents standardized tests from being as valid for them as for members of the majority culture. This contention is being examined by comparing regression equations predicting college performance from test scores and high school grades for Chicano and non-Chicano students in three California state colleges.

PROGRESS: Test scores and high school and college grades were recorded from the records of the three participating colleges. Test scores, both ACT and SAT, were only weakly related to grades, with a median correlation between .20 and .30, while high school grades showed a stronger relationship, producing a multiple correlation of test scores and high school grades with college grades generally between .40 and .50. Regression equations rarely differed between Chicanos and non-Chicanos, although sample sizes were small, on the order of 50 to 100 students, making comparisons questionable.

SAT Speededness Study (CEEB, 142-15) F. Evans

PURPOSE: The purpose of the study is to determine the relationship of the aptitude scores obtained under speed and power conditions and the racial and/or ethnic identity of the examinee. The reliability and validity, as well as the level of scores obtained under the various conditions, will be studied.

PROGRESS: Special tests have been administered to volunteer high school juniors in several high schools throughout the United States.

Developing 1974 Norms for the Preliminary Scholastic Aptitude Test (CEEB, 142-19)

W. B. Schrader
R. Jackson

PURPOSE: PSAT Norms were developed in 1960 and 1966. This study is designed to provide current norms for program use. It is also expected to throw some light on a possible trend in the performance of eleventh grade students in recent years. The decline of SAT scores has stimulated a great deal of interest and concern about this possible trend.

PROGRESS: Two hundred high schools have been selected at random from the comprehensive list of high schools maintained by ETS. Each school in the sample has been invited to test all of its eleventh-grade students at the time of the regular PSAT administration in October 1974. It is planned to replace any school which is unable to participate by inviting another school from the same region and similar in size to participate.

Evaluation of Preschool for Urban Children Project

(Minneapolis Board of Education, 268-6)

D. Norton

PURPOSE: The Minneapolis Board of Education, under a Title III grant, is developing a program to examine ways to enhance the gains made by children viewing Sesame Street. The evaluation efforts are to assess the effects of three different viewing conditions, each representing a different level of parental involvement. In one project component children attend the viewing center with their parents under teacher supervision. Another component consists of children viewing Sesame Street daily at the centers but without their parents. The primary means of communication for these parents is through a social worker and aide. The third component has children and parents viewing Sesame Street in the home. Parents attend weekly meetings aimed to develop a "parental teaching style." Children in all three groups receive support services including follow-up materials, field trip activities, and health services.

PROGRESS: The second operational year (1972-73) continues the Preschool for Urban Children in the same experimental framework as the first year. Pre- and posttests have been administered, and extensive data have been collected monitoring various aspects of parent and child attendance and viewing patterns, project management activities, as well as project health, instructional, and follow-up services. An Interim Evaluation Report was completed in January of 1973. A final report has been issued.

MAJOR FINDINGS: A battery of five Sesame Street Cognitive Tests, divided into 25 subtests, were the major assessment instruments used for the evaluation of child achievement. Evaluation results reported in the first year final report (1971-72) indicated that the major product goal for child cognitive achievement was met (with the exception of two minor subtests).

Minneapolis Pyramids Sibling Study, Phase III D. Norton
(Minneapolis Board of Education, 268-8)

PURPOSE: The Pyramids Reading Project is a major Title I ESEA effort in the Minneapolis Public Schools. Its impact should appear in reading scores of pupils but is hypothesized to be cumulative and not fully assessable for four or more years. The methodology of a sibling study promises better validity for project reports than other evaluative options. A longitudinal study of program outcomes is to be initiated.

PROGRESS: The project was initiated in June 1971. Outcomes of the first phases of the study reveal that older siblings did less well when tested at entry to first grade than did the pupils who received benefits from the Pyramids Reading Program in kindergarten during the 1969-70 school year; the differences that were noted paralleled those found in longitudinal city-wide test data. Phase II for 1971-72 is currently under way.

MAJOR FINDINGS: Initial findings were reported at the 1973 meetings of the American Educational Research Association. The report may be obtained through the ERIC system under the authorship of Daniel P. Norton and Richard W. Faunce.

CIRCUS (281-2) S. Anderson

PURPOSE: To develop a comprehensive assessment program for use in nursery school and kindergarten to diagnose the educational needs of individual pupils and to evaluate program effectiveness. The development process will include a national survey of the characteristics of children 4 1/2 to 5 1/2 and the preprimary programs they attend, with special attention to interrelationships among pupil and program variables.

PROGRESS: Based on experience in the ETS-OCD Longitudinal Study of Disadvantaged Children, the evaluation of Sesame Street, and other child development research, 15 measures of children's interests, cognitive and perceptual skills, information processing, problem solving, and divergent production have been developed, along with instruments to assess teacher and program characteristics. These are all designed to be usable by and meaningful to teachers and others without special training in psychology or research. After revisions of the measures on the basis of pilot administrations, two surveys of preschool children and programs were undertaken, one in kindergartens and the other in nursery schools. Representative national samples were obtained, using elementary school sending districts as the primary sampling units. The data collected in these two national administrations are now being analyzed.

Norming and Validity Studies of North Carolina Nursing Equivalency Examinations (Kate B. Reynolds Health Care Trust, 347-3)

W. B. Schrader
P. E. Vitella

PURPOSE: To collect and analyze statistical information to aid test users in interpreting performance on the North Carolina Nursing Equivalency Examinations. The primary objective of the study is to provide a basis for enabling Licensed Practical Nurses who have learned on the job to obtain credit-by-examination toward Associate Degrees in Nursing in North Carolina.

PROGRESS: The four tests included in the equivalency examinations were developed by ETS prior to the initiation of this study. Arrangements have been made to administer the tests to Associate Degree Nursing Students in 18 institutions in North Carolina at or near the end of the first year of the two-year program. Rating scales of competence have been developed with the advice of three persons familiar with A.D.N. programs in North Carolina. Each institution which is participating in the testing has been asked to supply data both on ratings and on first-year course grades.

Handbook for Conducting Task Analyses and Developing Criterion-Referenced Tests of Language Skills (Defense Language Institute, DLIEL-D, 474-1 through 474-7) J. Fremer

PURPOSE: To provide "state-of-the-art" information on criterion-referenced testing of language skills and to develop and try out appropriate test construction and analysis procedures that could be used by Defense Language Institute (DLI) - English Language Branch test development staff. The DLI prepares tests of English language skills that are used to determine whether nonnative speakers of English have attained enough English Language competency to receive technical instruction at military training schools. The language of instruction is English, and instruction is focused primarily on the native speakers who make up the majority of the classes. A nonnative speaker with insufficient command of English cannot keep up with other students.

PROGRESS: After a series of familiarization visits to language training and technical training sites, two technical courses, Air Traffic Control and Electronic Principles, were selected for intensive study. A literature review and survey of current practices were used as background for planning the collection and analysis of samples of the language used in instruction. Appropriate textbooks and other reading materials were gathered and observers/recorders used to obtain samples of classroom spoken language.

After the language samples were coded, they were analyzed to create both word frequency lists and summaries of such linguistic features as sentence length, tense, degree of embedding, etc. Test questions were written, using actual classroom language samples as stimuli, to tap aspects of language skill shown to be essential to comprehension of written and spoken language. Reading and listening comprehension pretests were developed and administered to students at various stages of language and technical training. Data from an existing language comprehension test and ratings of language proficiency by teachers were obtained for the students. In an initial testing phase, 51 students were tested and data obtained that were used to refine testing procedures. About 160 students participated in the final testing. Since these students represented a variety of subgroups, samples for any one analysis were small for drawing inferences about the populations of interest, but large enough to permit a demonstration of validation procedures for the criterion-referenced tests of language skills. A description of each of the steps in the project is contained in PR-74-11.

The test construction and analysis procedures developed during the project have been organized into a Handbook for use by DLI staff (PR-72-12). The handbook reviews the philosophy of criterion-referenced testing; outlines the development of a criterion-referenced test--from planning stage through score interpretation; and includes a review of the literature, annotated bibliography, glossary, and a section on the future of criterion-referenced measurement.

A Study of Empirical Option Weighting in the LSAT R. Reilly
(Law School Admission Council, 480-1)

PURPOSE: Differential weights for item alternatives were used to increase internal consistency of the LSAT. The scores generated with differential weights were then compared with rights-only scores in terms of their correlation with first-year law school grades.

MAJOR FINDINGS: Despite the large increases in internal consistency reliability, the validity for scores generated with differential weights was lower than the validity of rights-only scores. A Research Bulletin will be issued shortly.

Extension of the Study of the Possibility of Curvilinearity in Law School Validity Study Data (Law School Admission Council, 480-57)

W. B. Schrader
B. Pitcher

PURPOSE: To investigate the effect of normalizing predicted and actual law school grades on curvilinearity of regression.

PROGRESS: An exploratory study completed last year showed a statistically significant tendency for actual and predicted grades to differ systematically at the extremes of the predicted grade distribution. In this study, data for the same 30 student groups were reanalysed after normalizing both predicted and actual grades. The new analysis also used a more detailed comparison of predicted and actual grades than the exploratory study. In addition, the present study included a study of three more recent groups including a substantial number of black students. For these groups, separate analyses were made for the combined black and white student group and for white students separately.

MAJOR FINDINGS: It was found that normalizing had very little effect on the evidence of curvilinearity in the 30 groups used in the earlier study. However, the revised method of analysis did not confirm the previous finding of a systematic difference between predicted and actual grades. When validity coefficients for normalized variables were compared with validity coefficients for a linear transformation of the original variables, the normalized variables had a higher validity in 7 comparisons, and a lower validity in 15 comparisons, with 8 ties.

For the three law schools having a substantial number of black students, students in the top tenth on prediction earned grades significantly below prediction in one of six comparisons. This result occurred whether or not scores were normalized. Students in the low tenth on predicted grades showed a statistically significant tendency to earn grades below prediction in all three comparisons based on black and white students combined when the variables were normalized. None of the other comparisons for students in the low tenth was significant.

Prediction of Law School Grades for Mexican American and Black American Students (Law School Admission Council, 480-58)

W. B. Schrader
B. Pitcher

PURPOSE: To study differences in prediction equations for Black American or Mexican American students and for nonminority students in the same law school.

PROGRESS: Data for Mexican American and nonminority students were supplied by three law schools and for Black American and nonminority students were supplied by seven law schools. In one law school, separate analyses were done for 1971 and 1972 entrants, because the grading system differed for the two years. Analyses of covariance were performed for LSAT scores, Writing Ability (WA) scores, and undergraduate grades (UGPA) separately and for two combinations of scores--UGPA and LSAT, and UGPA, LSAT, and WA. In six law schools for which grades from two successive years of entrance were combined, the proportion of minority students was made equal for both years of entrance by removing nonminority students at random from the year having a smaller proportion of minority students. In addition to the analyses of covariance, validity coefficients were determined. Finally, predicted and actual grades were determined for minority and nonminority students using a simple equation based on data for both groups.

MAJOR FINDINGS: For both Mexican American and Black American students, the majority of analyses of covariance showed the same basic pattern: no significant difference in regression slopes, and a significant difference in intercepts. In each instance, predicted grades corresponding to a designated score on the predictor or predictors were higher for nonminority than for minority students. In three groups for which the sample size justified the use of Gulliksen-Wilks analysis of covariance, all five analyses in one group showed a significant difference in errors of estimate, a result which presumably arose because the standard deviation of law school grades was substantially larger for the minority students (Mexican American) than for the nonminority students. Eight of 55 analyses of covariance showed a significant difference in slopes. No systematic tendency for slopes to be larger for minority or nonminority students was discernible. In two of 55 analyses of covariance, no significant difference was found. From a statistical viewpoint, the findings indicate that in all but these two instances, separate equations would be warranted. The practical usefulness of this finding is seriously limited, however, by the fact that using separate regression equations amounts to using minority group membership as a predictor and by the fact that in 40 of 55 analyses for which slopes were equal but intercepts differed, predicted grades were systematically higher for nonminority than for minority students.

When a single prediction equation based on both minority and nonminority students was calculated for each of the 5 predictors or predictive combinations for each of the 11 groups, it was uniformly found that the mean predicted grade was higher than the mean actual grade for minority students.

Validity Study of LSAT Part-Score Reporting B. Pitcher
(Law School Admission Council, 480-59)

PURPOSE: This study will investigate the validity for predicting first-year average grades in law school of several part-scores derived from the separate section scores of the Law School Admission Test. All analyses are to be done separately by law school. Included within this study is a comparison of the validity of the traditional predictors, undergraduate average, LSAT and Writing Ability scores, for male and female law students at schools that have at least 50 female students in their 1971 entering classes.

PROGRESS: A sample of 15 law schools has been selected from the group of 99 law schools that participated in the 1972-73 Law School Validity Study Service, and a tape containing the data for the sample has been prepared for use in this study. The work on this study has been completed and two reports are being distributed to the Law School Admission Council. The first, titled "LSAT part-score validity study," is based on data from all 15 schools. The second, "Predicting law school grades for female law students," is based on data from 8 schools that had at least 50 female students in their 1971 entering classes.

MAJOR FINDINGS: On the average, there was no improvement in validity from the use of three LSAT part-scores (Reading, Data Interpretation, Principles and Cases) in place of the single LSAT score either when undergraduate grades and Writing Ability scores were also included as predictors or when they were excluded. Validity coefficients based on the prediction of first-year average grades in law school from the traditional predictors tended to be higher for female students than for male students, in spite of the generally greater variability of the male student groups with which the female groups were compared. No consistent trend toward either underprediction or overprediction of law school grades for female students was found.

Student Subgroups Validity Study B. Pitcher
(Law School Admission Council, 480-60)

PURPOSE: This study is designed to compare the prediction of first-year average grades in law school from the traditional predictors, undergraduate average grades, Law-School Admission Test and Writing Ability scores, for various subgroups of students pooled across law schools.

PROGRESS: Data for 90 law schools have been extracted from the tapes used for the Law School Validity Study Service in 1972-73, first-year average grades have been converted to a common scale, and a tape for use in this study has been prepared. Preliminary counts indicate that there should be enough data for comparing groups based on sex, ethnic origin, age, and undergraduate major, although some of the planned subcategories may not be usable. The statistical analysis of the special subgroup of students whose records included more than one set of LSAT and WA scores is essentially completed. Analyses based on other subgroups are under way.

Development of Graduate Student Rating Scales
(GRE Board, 540-59)

A. Carlson
R. Reilly

PURPOSE: The purpose is to develop a set of rating scales for use in evaluating graduate student performance. Behaviorally anchored rating scales have been developed and will be tried out in a small number of graduate departments in chemistry, English, and psychology.

PROGRESS: Three departments are currently collecting data which will be analyzed over the summer. Data from several more departments will be collected in the fall.

GRE Bayesian Validity Study (GRE Board, 540-64) R. Boldt

PURPOSE: The purpose of this study is primarily to gain experience in the application of a Bayesian model in the validity study setting and secondarily to accomplish a graduate school validity study using GRE variables. The model used allows the pooling of data in a situation where criterion observations may be scattered over a variety of institutions where they are not necessarily on the same scale. The model is being proved against a least squares model which allows a partial pooling of data.

PROGRESS: Data were solicited from departments of psychology and economics that require the GRE and the Advanced Tests for admissions but insufficient data were received. Data from the College Board Validity Study Service have been used instead.

MAJOR FINDINGS: The least squares and Bayesian systems are almost exactly equivalent in cross-samples from a correlational point of view. However, average errors of estimate indicate a slight superiority of the Bayesian system. Both systems yielded at least some negative regression coefficients: validity in a cross-sample was increased by arbitrarily reversing the sign of these negative coefficients. The final report is now in review.

Graduate Performance by Critical Incident (GRE Board, 540-75) R. Reilly

PURPOSE: The purpose of this project is to better define the criteria by which graduate students are evaluated through the use of a list of Behavioral Descriptions supplied by graduate faculty. The secondary objective is to contrast the relative importance of different criteria across three discipline areas representing the broad spectrum of graduate disciplines. Faculty from the areas of Psychology, Chemistry, and English utilized the list of the Behavioral Descriptions to describe the performance of graduate students with whom they have had contact.

PROGRESS: A report is available (RB-74-2).

MAJOR FINDINGS: Factor analyses revealed a total of eight factors which were relatively consistent across the three disciplines. Behaviorally anchored rating scales are currently being developed based on the factors.

Word Associations of Black and White College Students (GRE Board, 540-77) J. Campbell

PURPOSE: This project was designed to compare word assocation responses for college students from four different kinds of colleges in three different geographic areas: (1) selective predominantly white colleges, (2) less selective predominantly white colleges, (3) selective predominantly black colleges, and (4) less selective predominantly black colleges. Stimulus words were chosen from those used in verbal analogy and antonym items.

PROGRESS: Responses on 200 words have been obtained from approximately 3,000 students. These responses have been tabulated by sex for the four college groups noted above. A report is being prepared. (Study being done in collaboration with Dr. Leon Belcher of Texas Southern University.)

Women Doctoral Recipients (GRE Board, 540-78) J. Centra

PURPOSE: The main purpose of this study is to describe the current status and professional development of a sample of women who have received a doctorate during the past 20 years or so. Trends in their experiences and backgrounds will be compared with a sample of male doctorates matched by a field study and graduating institution. A secondary purpose of the study is to obtain information on the views toward their own graduate training of both females and males.

PROGRESS: Over 5000 graduates from the years 1950, 1960, and 1968 were selected. Among the areas covered in the study were the so-called "professional socialization" process, patterns of interruptions and obstacles to career development, demographic characteristics, and reentry problems. The questionnaire was mailed out during the Spring of 1973, and 81% of those at the addresses available responded. A final report entitled "Women, men, and the doctorate" will be available in Fall 1974.

GRE-Quantitative Coaching Study (GRE Board, 540-83) F. Evans

PURPOSE: This study is designed to determine the susceptibility of the GRE-Q and its various parts to a program of short term instruction and to determine the relative effectiveness of that program in raising the GRE-Q scores of candidates according to their distinct cultural background and sex. The study will be carried out in four parts: (1) feasibility, (2) exploratory, (3) developmental, and (4) operational. The feasibility study was conducted to determine if sufficient student and faculty cooperation can be obtained. A decision to continue the study was made based on the results of the feasibility study. The exploratory study will aid in the design and construction of curriculum and materials. The developmental phase of the study involves the actual design and production of curriculum materials. In the operational phase the program of instruction will be offered on several campuses and the results of that program assessed.

PROGRESS: The feasibility, exploratory, and developmental phases of the project have been completed. A curriculum covering eight weeks was developed and offered to senior volunteers at 12 colleges in the fall of 1973.

The curriculum for the experimental subjects consists of eight two-hour sessions, of which the first and last are devoted to testing. The second session consists of a presentation of and practice with various general testing strategies (e.g., pacing, when and how to guess). The second session also includes a presentation of the uses of tests in the admission process, the legitimacy of admissions tests (validity, reliability), GRE test directions and answer sheets. In short the second session is devoted to test familiarization which we hope will alleviate the candidate's anxiety about the admissions process and particularly about the GRE tests.

The third session consists of a short (40-item) mathematics test to assess the effects of the second session. The test was then reviewed item by item as a learning exercise. The test is a scaled down version of the GRE-Q in that it covers similar items and content.

Sessions 4 through 7 cover the basic content that is deemed necessary for adequate performance on the GRE-Q as identified in the exploratory study. Each of these sessions and the content they cover is listed below.

Session 4 - basic number facts, operations with fractions, lowest common multiple, greatest common divisor.

Session 5 - averages, ratio and proportion, and percent.

Session 6 - basic linear algebra, geometry.

Session 7 - data interpretation.

The materials for each of these lessons consist of a teacher's guide, classroom exercise book, diagnostic test, classroom test, and a workbook for out-of-class practice. The curriculum for the control group covers the same content and uses the same materials as the experimental group's but in a different order. Specifically, session eight, the posttest, was inserted between sessions one and two for the control subjects.

Thirty-four classes were conducted on 12 campuses. Colleges were selected in order to maximize the participation of minority students. On five campuses four classes (two experimental and two control) were held while on the remaining seven campuses only two classes were held (one experimental and one control). Seventeen teachers were recruited from either local high school faculties or from the faculties of special programs at the college, e.g., Academic Foundations Dept. or Educational Assistance Programs. Sixteen of the teachers attended a one-day training session in August 1973. At that time the structure of the course and the materials were reviewed. Nine of the teachers were employees of the college at which they taught our special course. All but two of those nine had had previous high school mathematics teaching experience and all were involved in special educational programs for inadequately prepared undergraduates. Of the teachers 5 were women, 6 Black-American, 5 Mexican-American, and 1 an American born oriental.

It was anticipated that the 34 classes would result in 340 experimental subjects and 340 control subjects. Further it was anticipated that we would have 140 Black (80 in predominantly Black schools), 50 Mexican-American, and 150 Caucasian experimental subjects and equal numbers of control subjects.

Data have been received from 11 colleges and have been processed. The attrition rate of subjects in the study was high necessitating major changes in analysis plans.

The Role of Cognitive Style in Students' Academic Evolution from High School Through Graduate and Professional School (GRE Board, 540-86; NIMH, 737-3)

H. Witkin
P. Oltman
F. Friedman
D. Goodenough

PURPOSE: One aim of this study is to assess the role of cognitive styles in students' academic choices and performance in the college period, selection of elective courses and majors, shifts in major, performance in various subject matter areas, post-college aspirations. A second aim is to assess the role of students' cognitive styles in predicting the decision to go on to graduate or professional school; the kind of school favored, eligibility for graduate or professional school, specialty selected after admission, and performance in graduate or professional school. A third aim is to explore continuity in students' choices and performance, as a function of cognitive style, from the high school through the higher-education levels.

PROGRESS: A four-year longitudinal study, through the college years, of an entire class of 1600 of a large municipal college was completed in 1971. Cognitive-style data were obtained at admission for all students in the class and a more intensive study was made of smaller selected groups showing different kinds of cognitive patterns. College and high school transcripts were also obtained for the entire college class. Students who went on to graduate school have been identified, and the school they are attending and area of specialization determined. Students who have gone on to various professional schools, and the schools they are attending, have also been identified. Analyses of data from this longitudinal study are now in progress. Some of the findings have been cited in Research Bulletin 73-11.

Development of Provisional Criteria for the Study of Scientific Creativity (GRE Board, 540-87)

N. Frederiksen
W. Ward

PURPOSE: This project was undertaken to develop tests of scientific thinking that may be used, at least provisionally, as criterion measures and that may be useful in studies of the creative performance of scientists. The tests will be based on job samples or simulations of the work of a behavioral scientist and will attempt to measure abilities required of a scientist in conceiving, planning, and evaluating research studies and in interpreting research data. Reliability and validity will be investigated. If the tests prove to be satisfactory psychometrically, they should be useful in investigating the correlates of various aspects of creative performance and providing dependent variables in experiments on situational factors that might influence the behavior of a scientist.

PROGRESS: Tests have been developed with the following titles: Formulating Hypotheses, Evaluating Proposals, Measuring Constructs, and Solving Methodological Problems. These four tests, each with six problems, were administered as "Tests of Scientific Thinking" to candidates taking the GRE Advanced Psychology Test in October of 1973, using an item-sampling procedure. Usable data were obtained for more than 4,000 candidates. Scoring procedures have been developed and tried out, and the scoring is nearing completion. Analysis of the data will provide information on the psychometric properties of the tests (reliabilities and intercorrelations) and correlations with GRE test scores obtained from the same administration. This information is relevant to the construct and discriminant validity of the tests. Correlations will also be computed with questionnaire responses dealing with amount and kind of training in psychology.

A substudy conducted at the University of Washington shows that the average judged quality of the candidates' best response is higher when candidates are asked to provide several hypotheses to the Formulating Hypotheses Test than when they are asked to provide only one (best) hypothesis.

National Assessment of Educational Progress J. Fremer
(Education Commission of the States, 553-1 through 553-4)

PURPOSE: To assess the attainment of educational goals in selected subject-matter areas by a representative sample of 9-, 13-, 17-year-olds, and young adults, and to measure progress over time by periodic reassessments.

PROGRESS: ETS has been contributing to the National Assessment project since its preliminary planning phases in 1964-65. Our early work included explorations of the goals in nine subject-matter areas and a number of special studies. Feasibility studies were undertaken of a variety of exercise (test item) formats; these included some designed to demonstrate a skill and many which required unusual stimuli and response modes. Other studies centered around the development of "easy tasks" to determine if it were possible to construct valid measures of what 90% or more of a group has achieved. Still other studies looked for ways to avoid cultural, sex, SES, geographic, and other biases. ETS has developed exercises in six assessment areas: writing, literature, science, social studies, music and art. (Other organizations have developed materials for reading, mathematics, career and occupational development, and citizenship.)

The schedule for administering National Assessment Exercises is as follows:

Assessment Year		Subject Areas
01	1969-70	Writing, Citizenship, Science
02	1970-71	Reading, Literature
03	1971-72	Music, Social Studies
04	1972-73	Mathematics, Science
05	1973-74	Career and Occupational Development, Writing
06	1974-75	[Note: Plans for future National Assessments were revised in the fall of 1974 to take into account reduced funding. Either one or two subjects will be administered in future years depending on the funding level.]

During the period July 1, 1973 to June 30, 1974, ETS held an exercise writers' conference and developed and field tested exercises for the Literature assessment.

MAJOR FINDINGS: NAEP has written numerous brochures and articles as well as reports of the results of the assessments that are already completed. A list of publications can be obtained from the office of the National Assessment Information Service.

Colleagues as Raters of Classroom Instruction (572-72) J. Centra

PURPOSE: To investigate the relationship between peer ratings of teaching based on classroom visits and the ratings given by students.

PROGRESS: Peer ratings, based on two visits by each of three faculty members, were compared to the mean ratings of students in each class on 16 items dealing with teaching effectiveness.

MAJOR FINDINGS: Colleague ratings of teaching effectiveness based solely on classroom observation would probably not be reliable enough to use in making administrative decisions on teachers. Convergence or agreement between colleague and student ratings occurred for a third of the items. A final report is available as RB-74-18.

Preschool Teachers' Beliefs on Effective Teaching Techniques and Their Relationships to Pupil Characteristics

W. Emmerich

(OCD, 581-14)

PURPOSE: The Enhancement of Learning Inventory (ELI) was constructed to assess a teacher's beliefs about the effectiveness of a variety of methods for teaching each pupil in her classroom. The present study examined these beliefs in relation to pupil cognitive and background characteristics for 35 teachers of economically disadvantaged preschool children, most of whom were enrolled in Head Start.

MAJOR FINDINGS: Individual differences among pupils prior to preschool entry were related to teachers' beliefs about effectiveness of teaching techniques toward the end of the preschool year. For example, pupils whose mothers had more years of school were judged by teachers as generally more receptive to classroom learning. However, processes engaged by classroom learning (cognitive skills and style) were more strongly and consistently related to teacher beliefs than were the child's family-background characteristics. Pupils exhibiting relatively high levels of cognitive skill prior to preschool were judged by teachers generally as most capable of classroom learning, especially when pupils were given considerable autonomy in structuring their classroom learning experiences. Also, children with moderate response tempos were perceived as more receptive to teaching efforts than children with either fast or slow response tempos. The child's sex and age at preschool entry also influenced teacher judgments. Different patterns of teaching style and teacher individuation were related to the pupil's initial characteristics and in several instances also to cognitive growth during the preschool year. While there was evidence that teacher belief patterns influence certain aspects of pupil psycho-educational development, the findings generally indicate that pupil behavioral characteristics may have a greater impact upon teacher behavior than vice versa.

A report is available (PR-73-12).

Development of Sex-Role Constancy in Disadvantaged Children (OCD, 581-15, 779-7)

W. Emmerich

PURPOSE: This developmental study of sex-role constancy in economically disadvantaged children continues work initiated as part of the ETS-Head Start Longitudinal Study. It investigates changes in sex-role constancy from ages 4-7, with particular attention given to the cognitive and environmental determinants of such changes. The sex-role constancy task assesses whether the child is able to hold gender constant despite suggested superficial changes in the sex-typed appearance or behavior of the depicted person. The task taps a process commonly used as a marker of the transition from preoperational to concrete operational thought in the Piagetian sense, thus extending a Piagetian approach to the study of social development. Originally proposed by Kohlberg, this measure of gender constancy is just beginning to be used in research, and the present study may be the first to apply it to the study of sex-role development in economically disadvantaged children.

PROGRESS: At ages 4 and 5 most children in our sample (over 1,000) were unable to hold gender constant on most of the task's items, suggesting that gender constancy is a relatively late developmental achievement in this population. However, there was also surprising evidence that many of these young children held constancy on certain items, suggesting that preoperational children may hold gender constant on bases other than concrete operational thinking. These pseudo-constancies could represent transient clusterings based upon motivational and affective states, such as fear that acting like someone of the opposite sex leads to punishment.

A series of analyses is being conducted to test a theoretical model suggested by the above initial findings. Initial findings are found in the following reports.

(1) Emmerich, W. Socialization and sex-role development. In P. B. Baltes & K. W. Schaie (Eds.), Life-span developmental psychology. New York: Academic Press, 1973.

(2) Emmerich, W., & Goldman, K. S. Boy-Girl Identity Technical Report. In V. C. Shipman (Ed.), Technical Report Series. Princeton, N.J., 1972.

Personality Structure and Development in Headstart Children (OCD, 581-15, 779-8)

W. Emmerich

PURPOSE: As part of the Longitudinal Study of Disadvantaged Children and Their First School Experiences, this study has the following aims: (1) Evaluation of a multivariate observation procedure for deriving personal-social measures on young children during relatively unstructured class periods; (2) Discovery of the structure (organization) of personal-social behaviors in Headstart children and possible variations in structure associated with the child's sex, age at entry, and first versus second semester enrollment; (3) Examination and interpretation of mean differences on peronal-social constructs associated with the child's sex, age at entry, first versus second semester enrollment, and their interactions; (4) Using a personality typology based upon the structural analyses, stability and change in personal-social behaviors during the Headstart year will be evaluated. A model for predicting the qualitative directions of such change will be tested; (5) The classroom personal-social measures will be related to independently assessed measures of cognition, style, and personality in the child; (6) It is hypothesized that parent-child and teacher-child relationships will jointly influence the child's personal-social behaviors in the classroom. These hypotheses will be tested by examining main effects and interactions among maternal and teacher variables with regard to their impact upon the classroom personal-social child measures.

PROGRESS: All Headstart period data have been collected, and those bearing on aims 1-4 have been coded and are being analyzed. A report on aims 1-4 is found in "Disadvantaged Children and Their First School Experiences: Structure and Development of Personal-Social Behaviors in Preschool Settings" (PR-71-20). A report on aim 5 is found in "Preschool Personal-Social Behaviors: Relationships with Socioeconomic Status, Cognitive Skills, and Tempo" (PR-73-33).

ETS-Head Start Longitudinal Study of Disadvantaged Children and Their First School Experiences

V. Shipman

(OCD, 581-15, 581-23)

PURPOSE: The purposes of the Longitudinal Study are to identify the components of early education that are associated with the cognitive, personal, and social development of young children; to determine the environmental and background modifiers which influence these associations; and to describe and evaluate the interaction of these as part of the young child's development over a period of six years (from approximately age 3 1/2 through third grade). The study has been sponsored and supported by the Department of Health, Education and Welfare, Office of Child Development (Grant #H-8256).

PROGRESS: This past year the study completed its sixth and final year of data collection as originally conceived. With most study children in the third grade, data collection included extensive parent interviews, individual and group child testing, classroom observations, teacher ratings, and questionnaires completed by school staff and community agency personnel.

Coding and processing of collected data has also continued. Specifically, almost all child test data through Year 5 (1972-73) have been processed.

Future activities will evolve around the continued processing, analyzing and reporting of data. Within the next year, a number of study reports are expected to become available.

MAJOR FINDINGS: See PR-71-19, "Structure and Development of Cognitive Competencies and Styles Prior to School Entry," December 1971; PR-71-20, "Structure and Development of Personal-Social Behaviors in Preschool Settings," November 1971; PR-72-13, "Demographic Indexes of Socioeconomic Status and Maternal Behaviors and Attitudes," June 1972; PR-72-18, "Structural Stability and Change in the Test Performance of Urban Preschool Children," September 1972; PR-72-27, "Technical Report Series," December 1972; PR-73-8, "Summary Report of PR-72-18," March 1973; PR-73-12, "Preschool Teacher's Beliefs on Effective Teaching Techniques and Their Relationships to Pupil Characteristics," May 1973; PR-73-18, "Development of Self-Regulatory Behaviors," May 1973; PR-73-21, "Characteristics of Urban Preschool Centers: Analysis of the Preschool Center Inventory," June 1973; PR-73-33, "Preschool Personal-Social Behaviors: Relationships with Socioeconomic Status, Cognitive Skills, and Tempo," August 1973; PR-73-35, "Interim Report," August 1973; PR-73-42, "Correlates and Implications of Self-Regulatory Behaviors," August 1973.

Computer and Communication Technologies and Efficiency in Education (583-20)

D. Jamison

PURPOSE: The purpose of this project is to examine the economic implications of the wider range of choice that new technologies provide educational decision-makers. For which educational tasks do computer and communication technologies appear cost-effective, and for which do they not?

PROGRESS AND MAJOR FINDINGS: The project is now about one-third completed, and the National Science Foundation has been requested to provide continuation funding. The nature of the results available to date can be gathered from the list of report titles below.

(1) The effectiveness of alternative instructional media: A survey - RB-73-57 (D. Jamison, P. Suppes, S. Wells).

(2) The impact of varying levels of computer-assisted instruction on the academic performance of disadvantaged students - RB-74-20 (S. Wells, B. Whelchel, and D. Jamison).

(3) Television and efficiency in higher education (D. Jamison and K. Lumsden).

(4) Projections of teacher salary costs (H. Guiterrez and D. Jamison).

(5) The Open University: A survey and economic analysis - RB-74-26 (K. Lumsden and C. Ritchie).

(6) The cost of educational television in the Ivory Coast (S. Klees and D. Jamison).

Evaluation of the Virginia Beach Extended School Year Program (585-60)

D. Powers

PURPOSE: To determine the effects of a year-round education plan on the achievement and attitudes of the children involved in a three-year pilot program.

PROGRESS: Baseline achievement, aptitude, and attitude data were gathered for a sample of 4th grade students enrolled in four pilot schools and two control schools in the Virginia Beach Public School System. Additional data have been collected on the same students after the year-round plan has been in effect for one year. Comparable data will be collected at the end of the second year of the program. Teachers' ratings of the degree to which their objectives corresponded to the items contained in the achievement battery have been obtained and are being used in the evaluation. Cross-sectional data obtained from the state testing program are also being used to independently assess program effects.

Differences between pilot and control groups and differences among various pilot groups are being examined. The interaction of ability level and type of school schedule is also being investigated.

At the first grade level, data have been collected in order to determine the effects on readiness of the early beginning necessitated by the pilot program.

A report on the effects of the first year is being prepared.

Accountability Design for the Cinnaminson Township (New Jersey) School District (Title III, 585-70)

R. Murphy
D. Powers

PURPOSE: A comprehensive plan for accountability is being developed for the Cinnaminson Township School District in cooperation with and under the direction of a local advisory committee. The advisory committee consists of the superintendent of schools, one elementary school principal, several teachers, parents, and concerned members of the community. The design is based on the use of longitudinal information in assessing educational effectiveness. The design includes a diagnosis to identify the probable causes of effectiveness and the development and implementation of plans for improvement based on the assessment and diagnostic results.

PROGRESS: To date, a general explanation of a design utilizing already existing test scores in the district for assessment has been made to the administrative staff, principals, teachers, and the advisory committees. Materials for obtaining information on characteristics of the educational program in mathematics are under development. A proposal for continuing the project in the 1974-75 school year is being prepared.

Evaluation of the New Approach Method (NAM) (OCD, 589-10)

D. Powers

PURPOSE: To determine the effectiveness of an experimental program, the New Approach Method (NAM), in teaching preschool-aged children beginning reading. The project was supported by funds from the U. S. Office of Child Development.

PROGRESS: A battery of instruments (both cognitive and attitudinal) was assembled or adapted and administered by a community-based field staff to children enrolled in the NAM program in their own homes and at five Trenton, New Jersey day-care centers (as well as to a randomly selected control group at one of the centers). Questionnaires were developed and administered to the parents of the children involved in the program. Additional data regarding the program operations were also gathered.

MAJOR FINDINGS: Results of the evaluation are presented in a final report (PR-73-47). The report contains a description and history of the NAM program as well as the results of the tryout. The program appeared successful to various degrees in accomplishing many of its objectives. Parental response to the program was quite favorable. Suggestions are made for future research on the program.

Survey of Adult Reading Performance (NIE, 592-2) R. Murphy

PURPOSE: A set of real-life reading tasks has been used to assess the reading skills of American adults. The tasks were developed on the basis of information obtained in a previous national survey of adult reading activities. A project report (PR-72-24) by Amiel Sharon describes the results of the activities survey. The survey of adult reading performance provides difficulty levels on approximately 170 items for a national probability sample and for a number of selected subgroups.

PROGRESS: The reading tasks have been administered nationwide by Response Analysis Corporation to a sample of 7500 adults 16 years old or older. In addition to responses to the reading tasks, information such as age, income, and educational background has been collected on the sample.

MAJOR FINDINGS: The reading activities survey showed that American adults spend a considerable amount of their time in reading while working, traveling, and shopping and that they consider such reading activities important. The national performance survey gives the extent to which various groups of adults can read materials intended to faithfully reflect these important reading activities. A project report (PR-73-48) gives the results of the National Reading Performance Survey.

A Descriptive and Analytic Study of Compensatory Reading Programs (USOE, 593-1)

D. Trismen

PURPOSE: In July 1971 ETS was requested by the U. S. Office of Education to develop design and analysis plans for a study of compensatory reading programs in grades 2, 4, and 6 of the U. S. public schools. The overall purpose of the study is to describe and to document the effectiveness of the various reading programs that are offered under the heading of compensatory education.

PROGRESS: The planning activity for the study took place during the period July-December 1971 and is described in the ETS Final Report for Contract No. OEC-0-71-3715. The first phase of the actual study was carried out during the school year 1971-72. Results are available in a report entitled A Descriptive and Analytic Study of Compensatory Reading Programs, Phase I. This phase consisted of a questionnaire survey in Spring 1972 of reading programs (compensatory and noncompensatory) offered in grades 2, 4, and 6 of a national sample of over 700 public schools. Data were gathered descriptive of schools, teachers, students, and instructional practices. From the Phase I data were then chosen 250 schools that were invited to participate in Phase II: 100 schools with compensatory reading programs funded by Title I; 100 schools with compensatory reading programs funded by sources other than Title I; and 50 schools offering no compensatory reading programs at all. An additional 30 schools with programs judged to be particularly innovative were also included. Data collection consisted of administering the same questionnaires that were administered in Phase I; pretesting and posttesting all students in grades 2, 4, and 6 using a reading achievement battery and a measure of attitudes toward reading; keeping records of students' attendance at reading instruction; and, finally, in the innovative schools, conducting a series of classroom observations. The analysis of Phase II data will involve relating achievement and attitude gain scores to program type in an effort to assess the relative effectiveness of the various programs.

A summer program study was conducted during the summer of 1973. Phase III, conducted during the school year of 1973-74, consisted of a closer look at 29 of the schools included in Phases I and II judged to be most and least effective for each program type delineated (see Major Findings below). Each school was visited twice between January and May 1974. Staff were interviewed and classrooms observed. Characteristics of reading programs that have been found to be most and least effective in terms of the criteria of this study were documented in detail. These findings are currently being assessed.

An in-depth study of a small number of schools with particularly effective reading programs is being planned for the fall of 1974.

MAJOR FINDINGS: Analysis of the Phase I (descriptive) data has just been completed. Five major program types were identified from the data provided by principals and teachers in questionnaires. Detailed descriptions of the program types, the schools they occur in, and the students they are intended for will be found in the final report for Phase I.

Survey of Disparities in Educational Achievements R. Feldmesser
(Ford Foundation, 595-1)

PURPOSE: This is an effort to determine whether existing data indicate any trends in the disparities in educational achievements between white and minority-group students since the time the Equal Educational Opportunity Survey (EEOS) was conducted in 1965.

PROGRESS: The research literature and reports of school-district and statewide testing programs have been searched for measures of disparities comparable to those in the EEOS, and both these measures and the EEOS data have been recalculated where necessary to maximize comparability. The final report is available.

MAJOR FINDINGS: Insofar as comparable data are available, they suggest that disparities have increased since 1965, though the trend may have been reversed in the last two years. However, the data are not sufficient in either quantity or quality to permit trend statements to be made with confidence.

Optimal Choice Theory (599-20) D. Jamison

PURPOSE: The purpose of this project is to undertake theoretical research on the theory of optimal individual and social choice.

PROGRESS: It is an ongoing project, and publications can be expected to ensue from it from time to time.

Reports available:

(1) Semiorders and the theory of choice (D. Jamison and L. Lau), Econometrica, 1973.

(2) The probability of intransitive majority rule: An empirical study.

(3) The nature of equilibrium with semiordered preferences (D. Jamison and L. Lau), in preparation.

Evaluation of Program for Training Dentists in the Out-of-Hospital Care of the Handicapped
(Robert Wood Johnson Foundation, 673-2 and 673-3)

J. Campbell
R. Flaugher
B. Esser
J. Knapp

PURPOSE: The Robert Wood Johnson Foundation is sponsoring a program in 11 dental schools for training dentists in the out-of-hospital care of mentally and physically handicapped Americans. This training program extends over a four-year period from July 1974 to June 1978. The ETS activities directed toward evaluation of the program's effectiveness began in January 1974 and is expected to extend to December 1980.

PROGRESS: Base line data are being collected currently, consisting of a test of specific technical knowledge, a measure of confidence for treating the dental problems of the handicapped, and a survey of background variables. These are being administered to the 1974 graduates of the 11 dental schools, for comparison with the results from later years, after the new program has been in operation.

A Survey of Skills and Abilities Needed for Graduate Study in Business (Graduate Business Admissions Council, 693-59)

J. Campbell
P. Casserly

PURPOSE: This project was designed to investigate the cognitive abilities necessary for student success in graduate schools of business.

PROGRESS: Job analysis type interviews were held with students, faculty members and administrators at 19 graduate schools of business. In addition to activities drawing on cognitive abilities, the interviews discussed affective propensities of students, their undergraduate preparation and work experience, and a number of other factors.

MAJOR FINDINGS: (1) The verbal and quantitative areas presently measured by the ATGSB are essential, (2) Lack of knowledge in mathematics causes difficulty for students in many schools, (3) Ability to write clearly and simply is an essential skill, (4) Some students who are low on "tolerance for ambiguity" have difficulty with business problems where some important factors cannot be reduced to known, quantifiable variables, (5) Other cognitive factors including general reasoning, inductive reasoning, and deductive reasoning are undoubtedly important, but appear to be adequately represented in the present composition of the ATGSB, (6) Motivation is of predominant importance. Several possibilities for assisting the schools to obtain better assessment of students' motivation should be explored. A final report has been issued (PR-73-44 and ATGSB Brief #9).

The Determinants of Two-Year-Old Children's Responses to Brief Separations from the Mother

M. Lewis
M. Weinraub

(NIH, 694-1)

PURPOSE: This study was designed to investigate the socio-emotional and cognitive correlates of children's responses to brief separation from their mother. Some children cry when their mother leaves them alone; others are hardly phased by her departure and continue playing happily. Surprisingly, there is very little evidence as to whether distress on maternal separation is indicative of a secure and healthy mother-child relationship or an insecure and unstable relationship. Evidence regarding the cognitive bases of the separation response is also sparse.

PROGRESS AND MAJOR FINDINGS: Fifty mothers and their two-year-old children were individually observed in a natural waiting room-like situation. After 15 minutes, the mother was signaled to leave the room: She left the room in whatever way she felt comfortable. The child was left alone in the room for 3 minutes (less if he became upset) before the mother returned.

There were large individual differences in children's responses to maternal separation: crying was observed in nearly one-half the children observed; playing was observed in fully one-half the children. Though there were no overall sex or social class differences in the number of children who cried or played during separation, there was some suggestion that more lower class boys cried than other children. Lower class boys were more likely than other children to wait at the door for their mother. Boys, in general, made more attempts to regain contact with their mother than girls. Interestingly, the only significant social class difference in the mean amounts of behavior during separation was that middle class children often verbally instructed their mother to return; no lower class children did so.

Large individual differences were observed in the mothers' styles of leaving the child. Some mothers slipped out of the room unobtrusively, some told their children they were leaving, and some mothers told their children what to do in their absence. The mother's style of departure was highly correlated with the child's response to her absence. Children whose mothers tried to slip out unobtrusively took longest to begin playing during separation. Children whose mothers told them of the departure took less time to begin playing, while children whose mothers told them of the departure and also told them what to do during the separation began to play the soonest.

Analysis of the data is still in progress. Measures of the mother-child relationship during free play, the mother's cognitive style, and the child's level of cognitive development will be used to predict the child's response to separation and the mother's style of departure. Sex and social class differences are being analyzed. A system is being developed to analyze sequentially the mother-child interaction during the event-filled departure situation.

A paper based on portions of this study was presented at the Eastern Psychological Association meetings, Philadelphia, April 1974.

The Effect of Stress on Dream Affect (NIMH, 737-3)

D. Goodenough
H. Witkin

PURPOSE: This study had two major goals: (1) to study systematically the consequences of viewing stress films on the emotional content of dreams and (2) to test the hypothesis that respiratory correlates of emotion are similar in waking and dreaming states.

PROGRESS: A manuscript describing the study has been submitted for publication.

MAJOR FINDINGS: Anxiety in dream reports was significantly increased on nights following stress film viewing. Emotion-respiration relationships appeared to be congruent across waking and sleeping states.

Genetic Factors in the Development of Field Dependence and Spatial Abilities: Intercorrelations Among Family Members (NIMH, 737-3)

D. Goodenough
H. Witkin

PURPOSE: It has been proposed that certain cognitive abilities have, as one determinant, a gene located on the X chromosome. If true, then one might expect mother-son and father-daughter correlations of about equal magnitude; lower father-son correlations; and mother-daughter correlations somewhere between these two values. Previous studies have confirmed these deductions for tests of visualization, but are contradictory for measures of field dependence. An attempt is being made to replicate and extend this work.

PROGRESS: Data analyses have been completed.

MAJOR FINDINGS: The pattern of family intercorrelations expected on the basis of the X-chromosome hypothesis was not found.

Genetic Factors in the Development of Field Dependence and Spatial Abilities: X-Linkage Study (NIMH, 737-3)

D. Goodenough
H. Witkin
I. Olkin

PURPOSE: It has been proposed that certain cognitive abilities have, as one determinant, a gene located on the X chromosome. If true, then in patterns of family transmission one might expect these abilities to show linkage with characteristics that are known to be determined by genes on the X chromosome (marker genes). An attempt is being made to test this hypothesis for a number of cognitive factors including field dependence, space, and visualization.

PROGRESS: Analyses of the data are almost finished. A report of the findings is in preparation.

Respiratory Correlates of Emotional States

(NIMH, 737-3)

D. Goodenough
P. Oltman

PURPOSE: Respiratory rate (or breath time) is not very clearly related to emotional states. It has been proposed that the breath cycle is composed of three relatively independent components (inspiration, expiration, and post-expiratory pause), and that a clearer understanding of the respiratory correlates of emotion might be had if these components are examined separately. Many hand scoring attempts to measure these components have been made in the history of research in this area, but these attempts have been largely unsuccessful because of the difficulties in reliably defining the transitions between components. An attempt is being made to solve this problem by developing a computer program which derives component times from curves fit to respiratory data.

PROGRESS: A report describing these studies has been submitted for publication.

MAJOR FINDINGS: The results show no effect of the stress films on total breath times and no correlation between breath times and subjective ratings of mood. However, expiration times significantly increase and pause times significantly decrease during the stress films.

Short-Term Memory Storage During Sleep
(NIMH, 737-3)

D. Goodenough
P. Oltman

PURPOSE: The literature suggests that learning during sleep is difficult if at all possible. The precise reason for this difficulty is unknown, however. An attempt is being made to test the hypothesis that auditory perception and short-term memory are intact during sleep, but the capacity to transfer information to long-term storage in easily retrievable form is impaired.

PROGRESS: Effective methods of stimulus presentation and retrieval have been worked out in pilot research.

Analytical Ability in Verbal Functioning (NIMH, 737-3)

P. Oltman
H. Witkin

PURPOSE: It has been repeatedly found that measures of the cognitive style dimension of field dependence-independence are related at low and mostly nonsignificant levels to vocabulary and other related verbal abilities. However, it seems possible that there do exist some aspects of verbal functioning, especially those requiring breaking up of contexts, which would relate to field dependence.

PROGRESS: In order to explore this possibility, 18 verbal tasks of varying types, including word knowledge, fluency, context-breaking, etc., were administered to a large group of college students, along with two traditional measures of field dependence. In addition, written and spoken free-response verbal samples were obtained from a smaller group, all selected to be high on verbal ability, and subdivided into field-dependent and field-independent subgroups. The resulting data will be analyzed in an attempt to uncover a cluster of verbal tasks which share common features with the field-dependence measures.

MAJOR FINDINGS: The scores for the 18 verbal tasks were intercorrelated and three factors were rotated by the varimax method. The first two corresponded to Guilford's symbolic and semantic content categories, while the third was a field-dependence factor. One verbal disembedding task appeared primarily on this third factor, in the male sample only.

Cognitive Style as a Factor in Dyadic Interaction Under Conditions of Initial Conflict (NIMH, 737-3)

P. Oltman
D. Goodenough
H. Witkin

PURPOSE: Previous studies have shown that participants in an interaction (teacher-student, patient-therapist, peers) who are similar in cognitive style are more likely to form positive impressions of each other, even after relatively brief periods of time together, than participants with contrasting cognitive styles. In the studies which yielded this result, the underlying social ground rule was one of working together toward a common goal (to have the student learn, to make the patient feel better). It is reasonable to expect, however, that other ground rules might lead to other outcomes with regard to mutual attraction, reflecting the role of situational factors as moderator variables in influencing the effects of match or mismatch. In the present study the social ground rule was one of initial conflict between the partners in the dyadic interaction with the requirement that the conflict be resolved. Under these conditions it was predicted that, because relatively field-dependent persons are more accommodating to the point-of-view of another and less likely to express or even experience opposition or hostility to others, dyads with at least one field-dependent partner would be more likely to arrive at a resolution of the conflict and, therefore, end up with positive feelings toward each other than dyads with two field-independent partners.

PROGRESS: Each of 40 subjects was required to spend a period of time in discussion, on one occasion, with someone similar in cognitive style, and on another occasion with someone dissimilar. The task was to work out a basis for agreement about an issue on which, in pretesting, they were found to disagree. All sessions were videotaped, allowing analyses of both verbal and nonverbal communication processes.

MAJOR FINDINGS: Data analyses already completed indicate that, as expected, dyads consisting of two field-independent partners significantly more often ended their discussion in disagreement and showed significantly greater negative feelings toward each other (as expressed in a post-interaction questionnaire) than dyads consisting of a field-dependent and field-independent partner or two field-dependent partners. Opinion changes, involving acceptance of the partner's position, were more frequent among field-dependent than field-independent partners. A report on these first findings is now being prepared for publication.

Laterality Study (NIMH, 737-3) P. Oltman

PURPOSE: Increasing lateralization of functioning ("leftness vs. rightness") in various response modalities has been observed over the course of development. Whether or not various types of lateralization (hand, eye, visual field, ear, etc.) are related remains an unanswered question. In order to explore this domain, a battery of tasks known to show lateral differences will be administered and intercorrelations computed. Some preliminary results suggest that extent of lateralization is related to the field dependence-independence cognitive style, relations will also be examined between the lateral differences on each of the tasks and field dependence. It may be, for example, that there exist several unrelated clusters of laterality tasks and that cognitive style is related to certain of the clusters and not others.

PROGRESS: A battery of cognitive style and behavioral laterality tasks has been administered to a group of subjects, along with EEG recordings of right- and left-hemisphere activity. Data analyses are in progress.

Studies of Young Adults with Uneven Cognitive Development (NIMH, 737-3)

P. Oltman
H. Witkin

PURPOSE: Factor analytic studies of the Wechsler Intelligence Scales have revealed that the subtest scores fall into three main factors: verbal comprehension, analytic, and attention-concentration. This study focuses on the first two of these. The first factor involves what is usually meant by verbal ability: vocabulary, information, and comprehension. The analytic factor is actually quite similar to the field dependence-independence cognitive style. The fact that verbal and analytic or field-dependence tasks are generally uncorrelated implies that while some individuals would be relatively consistent in normative level on these two factors, others would be discrepant in varying degrees. To study some of the implications of evenness vs. unevenness in level of verbal and analytic performance, various groups can be composed, some of which show equally high, medium, or low performance in the two areas, and others which show discrepancies favoring either verbal or analytic performance.

PROGRESS: A group of college students with verbal ability within the upper quartile has been selected, and further selection within this group has yielded a subgroup equally high on analytic ability and a second, discrepant group, which is in the lower quartile on analytic performance. A wide variety of verbal behavior and other data has been sampled in these groups, including structured and free-response verbal tasks, interviews, vocational interest data, and college transcripts. Intensive study of these sources of data will provide evidence on the impact of verbal-analytic discrepancy on verbal functioning, on patterns of interests, and on various aspects of the individual's academic career.

Cognitive Patterning in Sighted Children with Bilateral Retinoblastoma (NIMH, 737-3)

H. Witkin
F. Friedman

PURPOSE: In previous research we found that children totally and adventitiously blind from retinoblastoma were markedly superior in analytical ability (that is, they were field independent), but not in verbal-comprehension ability or attention-concentration ability. The frequently cited cognitive superiority of children with retinoblastoma thus seems specific to the analytical domain. This phenomenon is of particular interest against the fact that retinoblastoma is caused by a specific autosomal gene defect, although at this moment the findings obtained may as well be explained on an experiential basis. This study is the first in a series intended to pursue further the possible roles of genetic and learning factors in the cognitive patterns observed in retinoblastoma children.

PROGRESS: The subjects of this study are children with bilateral retinoblastoma, in whom vision was saved, so that while having the gene defect, they also had the same opportunities for cognitive learning as normal sighted children. Such cases are quite rare and the search for them is being continued. As each child is located he is assessed on an extensive battery of cognitive tests. (Study being conducted in collaboration with the Columbia Presbyterian Eye Institute.)

Cross-cultural Studies of the Development of Psychological Differentiation (NIMH, 737-3) H. Witkin

PURPOSE: The purpose of this study is to assess the role of ecological, cultural and socialization factors in the development of psychological differentiation.

PROGRESS: Recent cross-cultural studies of psychological differentiation have suggested that mobile hunting groups, living in homogeneous surrounds, are likely to have loose, limited social organization, to emphasize autonomy in child-rearing and to be field independent in cognitive style. In contrast, sedentary agricultural groups, living in a heterogeneous surround, are likely to have a hierarchically organized society, to emphasize conformity in child-rearing and to be field dependent. Each cluster may be conceived as adaptive to its ecological context. Active planning is now in progress on a study to be started in the field early next year which will put this "ecological hypothesis" to further test by comparing Pygmy hunters living in deep forest with Black farmers living in clearings in the same forest in the Central African Republic. Assessment will be made of cognitive style and direct observations conducted of parent-child interactions. Sex differences in cognitive style, in relation to socialization practices followed with boys and girls, will also be studied in the two cultures. Finally, a study will be made of acculturated Pygmies who have grown up in the sedentary farming setting, and who have had no engagement with the forest hunting ecology.

Modes of Representation of Presleep Experiences in Dreams (NIMH, 737- 3)

H. Witkin
D. Goodenough

PURPOSE: This study seeks to investigate the processes by which waking thoughts, images and feelings find transformed representation (in symbols, metaphors, etc.) in the language of the dream. Of particular interest is the relation between cognitive style and individual differences in mode of representation of presleep experiences in dreams.

PROGRESS: Data gathering for this study has already been completed, and data analyses are under way. The subjects, half of whom were field dependent and half field independent, came to sleep in the laboratory, where their EEG and eye movement were monitored, on five occasions, a week apart. The first of these was a practice session, and on two of the remaining four occasions subjects viewed a stress film immediately prior to sleep and on the other two occasions viewed a neutral film. Subjects were awakened from each rapid-eye-movement period and asked to report anything that had been going through their minds prior to awakening. The large body of dreams collected in this way, following upon the various presleep stimuli used, provide the data now being analyzed for exploring the relation between presleep experiences and dreams as a function of cognitive style.

Teacher-Student Interaction, Student Achievement and Mutual Attraction Between Teacher and Students as a Function of Match or Mismatch in Cognitive Style

(NIHM, 737-3)

H. Witkin
D. Goodenough
P. Oltman
C. A. Moore
F. McDonald

PURPOSE: One aim of this study was to confirm the finding of past investigations that teachers and students similar in cognitive style tend to evaluate each other more positively than teachers or students who are mismatched. Another purpose was to determine whether students learn more and are more motivated to learn when taught by teachers similar to them in cognitive style. Still another purpose was to identify the processes of interaction through which match or mismatch influence teachers' and students' feelings toward each other and student achievement. A further purpose was to identify features of teaching style and learning style which are a function of cognitive style. A final purpose was to identify the ways in which teachers, whatever their own cognitive style may be, adapt their teaching methods according to the cognitive style of the student being taught. It is anticipated that the data gained in pursuing the last purpose will contribute to the development of teaching methods specifically appropriate to students of different cognitive styles.

PROGRESS: The study conducted in this area had the additional objectives of examining the role of cognitive style in teacher behavior and student learning. Teachers selected as field dependent or field independent taught a minicourse to small groups of field-dependent and field-independent students. A curriculum was specially prepared for the course to allow expression, in both course content and teaching techniques, of likely preferences of field-dependent and field-independent teachers. Sessions were videotaped and audiotaped and, in addition, teacher behavior was coded during class sessions. At the end of the minicourse, measures were obtained of student achievement and interest and teacher-student interpersonal attraction; and the teacher was interviewed about her course preparation, her classroom experiences and her reactions to each of her students. The data from the study are now being analyzed.

Psychological Differentiation as Related to Antisocial Behavior, Impulsivity, and Endocrine Functioning in a Prison Population (NIMH, 738-1)

P. Oltman
H. Witkin
D. Owen

PURPOSE: Past studies have shown that psychopathology takes different forms in more differentiated and less differentiated persons when they suffer psychological breakdowns. Taking criminal behavior as also reflecting in some degree lapses in control, this study seeks to determine the forms criminal behavior assumes in more differentiated and less differentiated persons charged with criminal offenses. Also being examined in this study is the relation, in a group of men, between extent of field dependence and testosterone level. This relation is of particular interest against the repeated finding that not only do men tend to be more field dependent than women, but within the sex those who score in the more masculine end of masculinity-femininity scales tend to be more field independent. Finally the study is examining the relation between testosterone level and aggression.

PROGRESS: 195 prisoners from the psychiatric prison service of a large New York City municipal hospital were the subjects in this study. Using all records available on each prisoner (the arrest report, prison chart, probation report) classification of crimes and methods of carrying out crimes were made, according to several hypotheses regarding the relation, first between psychological differentiation and forms of antisocial behavior, and, second between aggression and endocrine functioning. Similarly, classifications have been made of biographical data gathered through questionnaires administered to these prisoners. Finally, a scheme for assessing nature and effectiveness of controls was developed and applied to the figure drawings made by the prisoners. These data, together with the results of a battery of tests of field dependence, which were given to all the prisoners, are now being analyzed.

Relationship of Height to Criteria of Antisocial Behavior (NIMH, 738-1)

D. Owen

PURPOSE: It has been suggested that taller adolescents are more likely to be identified as delinquent than their shorter peers. The possibility has also been raised that convicted criminals who are taller than average may receive longer sentences than similarly charged men who are shorter. It is important to determine the extent, if any, to which added height influences the criteria of antisocial behavior, such as arrest record, conviction, or length of sentence. Therefore, an attempt is being made to assess the relationship between height and the existence of an arrest record. Also, an attempt is being made to assess the extent to which similar criminal cases, after a guilty verdict has been obtained, are disposed of differently as a function of the height of the criminal defendant.

PROGRESS: From a starting population of 31,455 consecutive male births during 1944-1947 in Copenhagen, Denmark, the heights of 28,588 individuals still living have been obtained. All civil and penal violations charged against any of these men and the dispositions of their cases have been collected from the police. A coding scheme to accommodate these crimes and resulting criminal sanctions has been devised, and the data have been coded. A computer tape file for the criminality data has been created and data analyses are in progress.

Genetic Factors in Field Dependence and in Aggressive Behavior: Studies of Sex Chromosome Anomalies

(NIMH, 738-1)

H. Witkin
D. Goodenough
D. Owen

PURPOSE: This study is attempting to identify noninstitutionalized men with two kinds of sex chromosome aberrations, one with an extra Y chromosome (XYYs), the other with an extra X chromosome (XXYs). The data obtained with the groups thus identified, along with the data from comparison groups of normal XY men, will be used for two purposes: first, to check a model of the role of the sex chromosomes in field dependence, and second, to check the hypothesis recently put forward that an extra Y chromosome contributes to the development of aggressive behavior. Since groups institutionalized for antisocial behavior which are, of course, highly biased populations, have been used in almost all the work done on the aggression hypothesis to date, what is especially needed are studies of noninstitutionalized groups.

PROGRESS: A case-finding study has already been completed in Copenhagen, Denmark, where social records are particularly well suited both for identifying subject populations and for obtaining evidence on antisocial behavior. All males (31,380) born in Copenhagen between 1941 and 1945 were identified. Using height data available for these men at the draft boards, the group in the top 15% for height were selected for sex-chromosome determination. For 91.5% of the total tall group, it was possible to locate and complete sex-chromosome determination. Data were obtained from available social records for the sex-chromosome anomalies who were found and matched controls. Intensive individual case studies are now in progress of the probands and their controls.

MAJOR FINDINGS: The XYY and XXY groups of men tend to show an elevated crime rate. Their performance on an army intelligence selection test indicated that tall groups are, on the average, substantially lower in intellectual functioning than XY controls, although it will be necessary to wait on the findings from the individual case studies now in progress to determine whether the lower level of intellectual functioning of the probands is the result of a specific cognitive defect or of general cognitive inferiority. The data analyzed to this point suggest, as a working hypothesis, that the elevated crime rate found in XYYs, and apparently in XXYs, may be the result of some kind of cognitive defect. This hypothesis contrasts with the alternate hypothesis that an extra Y (or extra X) chromosome is directly responsible for heightened aggression.

Sex Chromosome Anomalies in a Male Prison Population
(NIMH, 738-1)

H. Witkin
P. Oltman
D. Owen

PURPOSE: The proposal has been made that men with an extra Y chromosome are overrepresented in prisoner groups, compared to the general population, suggesting that the extra Y chromosome is associated with aggressive behavior. At the moment this "aggression hypothesis" is hardly proven. The data on hand bearing on this problem have often come from studies inadequately controlled and suffering a number of selection biases.

PROGRESS: Nine hundred eighty-three consecutive admissions to the prison ward of the psychiatric service of a large New York City municipal hospital were karyotyped. The small number of XYYs (only three) identified in this group is not beyond the frequency to be expected in the general population. This finding is inconsistent with the aggression hypothesis. The difference between the XYY frequency in our group and the higher frequency of XYYs found in some other mental-penal groups that have been examined may be related to differences in admission policy. It is noteworthy that four XXYs were also identified in our prison group, again a frequency not different from chance. The similar frequency of XYYs and XXYs is consistent with the results of several other recent studies. A paper on the results of this study is now being prepared for publication.

Reference Measures for Cognitive and Personality Factors (ONR,739-1)

H. H. Harman

PURPOSE: Under the sponsorship of the Office of Naval Research, ETS is conducting research to identify and to present measures of well-established factors by updating the previous Kit of Reference Tests for Cognitive Factors and by extending it to the noncognitive domain.

PROGRESS: Work in the cognitive area has involved three phases: (1) development of divergent production tests comparable to those in the 1963 Kit for which ETS did not hold the copyright, and the field testing of them; (2) review and modification of other tests of the 1963 Kit to be included in the new Kit; and (3) development of new marker tests for factors that have been established in the literature since 1963, and field testing at the Navy Training Center in San Diego.

Work in the noncognitive area has consisted of a tryout of our personality items on some 4,000 recruits at the Navy Training Center in San Diego. Also, the scales were administered to a sample of approximately 250 men and women undergraduates at the University of Oregon who had already taken two other personality inventories. Thus it is possible to look at the structure of our scales in a broader context. Two reports are now in progress regarding these personality scales. As a final product in this area, plans are to develop a Guide to factor-referenced temperament scales. This will include references to scales developed by other researchers that fit into the domain of concern of the current project.

MAJOR FINDINGS: The following reports have been released:

Toward the Establishment of Noncognitive Factors Through Literature Search and Interpretation. J. W. French (PR-73-29)

Cognitive Factors: Some Recent Literature. R. B. Ekstrom (PR-73-30)

Proceedings of Symposium: Toward the Development of More Comprehensive Sets of Personality Measures. H. H. Harman (RM-73-29)

Psychometric Tests as Cognitive Tasks: A New "Structure of Intellect." J. B. Carroll (RB-74-16)

Problems of Replication of Seven Divergent Production Factors. R. B. Ekstrom, J. W. French and H. H. Harman (PR-74-14).

Perceptual Factors in Cognition and Personality

(ONR, 741-10)

S. Messick
J. French

PURPOSE: This study investigates factor analytically the generality in cognitive and personality domains of the perceptual factors of speed and flexibility of closure.

PROGRESS: Experimental tests of verbal and cognitive closure and flexibility in reasoning and problem solving were developed and included in a test battery with selected personality and reference ability measures. The test battery was administered to more than 500 cadets at the U. S. Naval School of Aviation Medicine, Pensacola, Florida. A factor analysis of intercorrelations among 35 scores on perceptual and cognitive tasks has produced about 14 factors, which were rotated to oblique simple structure. Twenty scores from personality measures were also projected into the space by extension methods. Examination of the results indicated factors interpretable in terms of speed and flexibility of perceptual closure, of verbal closure, and of cognitive closure. Intercorrelations among the primary factors were factor analyzed to produce four second-order factors. A hierarchical transformation was performed to place the first- and second-order factors in the same orthogonal framework and to obtain the loadings of the tests on the second-order factors. The second-order factors appear to reflect verbal comprehension, general reasoning, figural articulation (i.e., facility in structuring and differentiating figural stimuli), and symbolic articulation. A report, "Dimensions of Closure in Cognition and Personality," has been completed and accepted for publication in *Multivariate Behavioral Research*.

Disattenuated Partial Correlation F. Lord
(ONR, 746-3, 746-4, 772-21)

PURPOSE: To find a significance test for the hypothesis that the correlation between two variables disappears when the true score on a third variable is held constant.

PROGRESS: An "optimal asymptotic test" was found. Monte Carlo studies have been conducted in order to compare this significance test and tests based on the sample corrected and uncorrected partial correlation coefficients.

MAJOR FINDINGS: The development of the significance test and a numerical example are presented in RB-72-6, which has been accepted for publication in Educational and Psychological Measurement.

Estimating Item Parameters (ONR, 746-3, 746-4; NSF, 816-1) F. Lord
M. Wingersky

PURPOSE: To find a practical method to estimate the parameters of the characteristic curve of each item in a test.

PROGRESS: A method for estimating ability and item parameters using information provided by omitted responses has been developed, programmed, and tried out on a variety of tests. An improved method is being programmed.

MAJOR FINDINGS: "Estimation of Latent Ability and Item Parameters When There Are Omitted Responses" (RB-73-37) describes the method and gives theoretical and empirical justification for it. The program, documented in RM-73-2, is available for general use.

Testing if Two Measuring Procedures Measure the Same Dimension (ONR, 746-3, 746-4, 772-21)

F. Lord

A convenient statistical technique is described for testing the hypothesis that two sets of measurements differ only because of errors of measurement and because of differing origins and units of measurement. (RB-73-41)

Published in Psychological Bulletin, Spring 1973.

Dimensions of Perceived Item Similarity S. Messick
(USPHS, 752-2)

PURPOSE: To ascertain the number and nature of dimensions of perceived similarity among personality statements (in terms of similarities in likelihood of endorsement) in an effort to appraise the relative contributions of perceived content consistencies and perceived desirability consistencies in determining judgments of endorsement frequency.

PROGRESS: Twenty-three MMPI items were selected to span six dimensions previously obtained in analyses of MMPI responses. All possible pairs of these 23 statements were presented to 150 college students, who were asked to rate the similarity of the members of each pair with respect to their probability of endorsement. The 150 x 150 matrix of sums of cross-products among individuals was computed and factored according to the individual differences model of multidimensional scaling. Four dimensions of viewpoint about the similarity of these items appeared, and separate multidimensional spaces were derived for each viewpoint. These perceptual spaces were rotated orthogonally, and the dimensions were interpreted. One important distinction reflected in most of these spaces is the separation of desirable from undesirable item content. A report is in preparation. (Study done in collaboration with Dr. D. N. Jackson of the University of Western Ontario.)

Leveling-Sharpening as a Cognitive Style
(USPHS, 754-7)

F. Damarin
V. McKenna
S. Messick

PURPOSE: This research has three major purposes and several minor ones. The first purpose is the replication and extension of earlier findings by Gardner, Jackson, and Messick on the cognitive style of leveling-sharpening. The tests which are involved in this research include many psychophysical measures. A second purpose is to discover whether the leveling-sharpening trait or other psychological traits can serve as the parameters of the psychophysical law, as Damarin has suggested they should. Since many leveling-sharpening measures are similar in their formal structure both to measures of figural aftereffects and to measures of retroactive inhibition, a third purpose of the research is to discover whether "levelers" also show larger FAE's and more retroactive inhibition as McKenna has hypothesized.

PROGRESS: Forty-seven objective tests, including putative measures of leveling-sharpening, field independence, category width, neuroticism, figural aftereffects, retroactive inhibition, verbal satiation, rote memory, memory span, meaningful memory, and psychophysical judgment, and a variety of questionnaires, including most of the known response styles, were administered to 147 freshman men and women at the College of William and Mary in Williamsburg, Va.

MAJOR FINDINGS: In an earlier study, now completed, cognitive style measures showed an impressive ability to predict individual differences in the recall of photographs of human faces, thus supporting assertions to this effect by Witkin (cf. RB-63-21).

A later study developed evidence that several different types of leveling-sharpening may exist. One of these is reasonably similar to the dimension that Gardner and his associates interpreted as the tendency to assimilate perceptions to memories.

The same research yielded evidence that individual parameters of the psychophysical function may correlate over different types of judgments, giving some support to Damarin's "trait law" hypothesis. A more refined analysis of these data is now under way.

McKenna's factor analysis of memory tests suggests that there are memory styles associated with a preference for errors of omission vs. commission. A report on this work is available (RB-68-28).

Personality Organization in Cognitive and Social Processes (USPHS, 754-7) S. Messick

PURPOSE: This project attempted (a) to clarify in a series of multivariate studies the nature of several cognitive styles and to investigate their interrelations with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; (b) to analyze sex differences in cognitive and personality structure; (c) to assess the stability of cognitive styles and the correlates of reliable changes; and (d) to evaluate the role of cognitive styles in interpersonal and social functioning.

PROGRESS: (a) Categorizing Styles and Cognitive Structure (with Nathan Kogan). This study attempts to clarify two basic questions: (1) Is there a general dimension of breadth of categorization yielding individual consistencies in categorizing behavior across varied stimulus domains and measurement procedures; and (2) what are the cognitive, intellectual, and personality correlates of stylistic consistencies in categorization? Group-administered measures of verbal and quantitative category-width judgments, object sorting, cognitive complexity, preference for complexity, and extremity of judgment were administered to approximately 160 college students, along with several other cognitive, intellectual, and personality variables. Factor analyses were performed separately for male and female subjects. The results indicate that categorizing consistencies exist, but that they are multidimensional and their organization is somewhat different for the two sexes. Females appear to display more generality across categorizing tasks than do males: several dimensions relevant to categorizing emerged for both sexes, but they appeared to be somewhat more differentiated and task specific for males. Personality measures were found to be implicated with categorizing and judgmental consistencies in females, whereas they were more directly associated with intellectual abilities in males. Although the emergence of several categorizing factors in this study indicates the operation of differential stylistic tendencies in this domain, the factors were at the same time found to be correlated with each other, suggesting some generality at a second-order level. Factor analyses of the intercorrelations among the primary factors produced three second-order factors for males and two for females. In both sexes, one of the second-order factors reflected general categorizing consistencies that cut across the stimulus and process distinctions embodied in the various first-order factors.

(b) Cognitive Complexity: Dimensions of Role Constructs (with Nathan Kogan). In conjunction with the study of categorizing styles described above, an attempt was made to assess certain aspects of cognitive complexity using multidimensional scaling techniques. Much recent work on cognitive complexity has used a procedure developed by George Kelly for assessing personal constructs--the Role Construct Repertory Test (Rep Test), in which the subject is presented with several triads of figures significant in interpersonal relations, such as mother, self, and best friend. The subject is then asked to indicate which two members of the triad are most alike and to write down both the basis for their similarity and the reason why the third member is different. The number of different constructs listed has been used as a measure of cognitive complexity.

The 160 subjects in the categorizing study described above were administered a modified form of the Rep Test, and they were also asked to judge on a 10-point scale the degree of similarity or difference between all possible pairs of the 15 role figures. Cross-products matrices for these similarity ratings were computed separately for males and females and were analyzed according to the individual differences model of multidimensional scaling. Six dimensions of viewpoint about stimulus similarity were obtained for the females and seven for the males. The individual scores on these dimensions were then correlated with the cognitive and personality measures included in the categorizing study in an attempt to ascertain some of the properties of individuals holding each viewpoint. In this case, meaningful patterns of significant correlations were obtained between the viewpoint scores and such measures as acquiescence, impulsiveness, rigidity, unconventionality, preference for complexity, category width, and conceptual differentiation.

Two composite stimulus spaces were also derived, one to represent the similarity judgments of the males on the average and one to represent the similarity judgments of the females on the average. For the females, four dimensions were required to represent the average similarity judgments among 15 stimuli, whereas for the males three dimensions appeared adequate. The same three dimensions appeared in both spaces, but the fourth female dimension provided a distinction in sex-role relations (they described similarities between "mother" and "sister" on a separate dimension from similarities between "father" and "brother") that was not emphasized by the males. One report is available and another is in preparation.

(c) Cognitive Interference and Flexible Control. The term "constricted vs. flexible control" refers to consistent modes of reacting to interfering and contradictory cues. This cognitive control has been most frequently assessed in terms of interference scores on the Stroop Color-Word Test. In the present study, a group-administered version of the Stroop Color-Word Interference Test was developed, as well as three other group procedures designed to measure susceptibility to cognitive interference in the face of differing amounts of competing response and at various levels of cognitive involvement. An attempt was made in developing these procedures to vary the ease of separating competing cues in the interference conditions, so as to permit an empirical separation of two related processes thought to be important in withstanding cognitive interference effects: (1) facility in restricting attention to the relevant aspects of the stimulus and the response and (2) facility in actively inhibiting the competing response and other disrupting tendencies. The four tasks were administered to approximately 90 college males, along with the original individually-administered Stroop Color-Word procedure and several other cognitive and personality measures. Both control and interference sections of all five Stroop-type tests were found to correlate highly with each other (median r of .53), but specific processes involved in each task were differentiated in a factor analysis, as were the cognitive and personality correlates of the factors. A factor analysis of intercorrelations among nine first-order factors uncovered two second-order dimensions, which appeared to reflect (1) the ability to withstand disruption in serially organized tasks and (2) the ability to isolate and identify the appropriate aspects of the stimulus and response. A regressed interference score was also obtained for each of the five procedures to reflect that part of the interference performance which was independent of the control performance. These regressed scores possessed substantial reliability, and although they were not related to each other, they exhibited different patterns of significant correlations with personality measures.

(d) Leveling-Sharpening and Assimilation Tendencies (with Fred L. Damarin and Virgil McKenna). The cognitive style of leveling-sharpening has been postulated to account for consistent individual patterns of performance in psychophysical tasks and in the relative differentiation of free association responses. These cognitive consistencies have been attributed to individual differences in the activity of the memory process of assimilation. Levelers assimilate in the sense that they tend to merge perceived objects or stimulus events with similar but not identical objects recalled from previous experience.

In an effort to clarify possible multiple determinants of leveling-sharpening consistencies, the present study attempted to specify the concepts of assimilation and contrast in many different ways as a basis for developing a variety of new experimental tests.

One procedure that was highlighted finds its prototype in the Schematizing Test, which has been used as a criterion measure of leveling-sharpening in other laboratories. The Schematizing Test provides two scores: a lag score measures the subject's tendency to underestimate change in a series of squares that gradually increase in size, and an accuracy score measures his ability to estimate magnitudes accurately within a fixed range of stimulus values.

By also including tasks with different types of stimuli and different judgmental conditions, the present study focused upon the construct validity of alternative interpretations of leveling-sharpening and assimilation-contrast. A battery of several experimental psychophysical judgment tasks was administered to 150 college students, along with measures of field-independence, category width, figural aftereffects, retroactive inhibition, verbal satiation, memory abilities, and various personality characteristics. A factor analysis produced 17 factors, which were rotated to oblique simple structure. An analysis of the intercorrelations among the first-order factors produced four second-order dimensions. One of the factors could very plausibly be interpreted as a measure of assimilation or dedifferentiation of memory traces over time, but the lag score from the Schematizing Test was not related to it. Another factor appeared to represent uncriticalness in judgment, a general tendency to judge similar stimuli as the "same" as opposed to "different," but again the lag score was unrelated. However, the schematizing score was found to be associated with one of the second-order dimensions, which could be interpreted in terms of leveling-sharpening. One report has been completed (RB-68-28) and another is in preparation.

(e) Tolerance for Unrealistic Experiences (with Norman Cliff). This dimension, also called tolerance for instability, refers to the differential willingness to accept perceptions at variance with conventional experience. The concept was introduced originally to account for data linking form-boundedness and form-lability on the Rorschach with responses to unusual kinds of perceptual experiences, such as those arising in apparent movement and aniseikonic distortion.

The present study attempted to appraise the reliability of several measures of apparent movement and aniseikonic distortion, and to evaluate the intercorrelations among them for evidence of stylistic consistencies. Measures of aniseikonic distortion and perceived range of apparent movement were obtained from a sample of 50 undergraduates. Moderate reliabilities were obtained for certain measures of aniseikonic distortion. High reliabilities were achieved for movement threshold and simultaneity threshold scores in the apparent movement procedure, but the two thresholds correlated positively with each other (.46). Correlations between aniseikonic measures and the movement and simultaneity measures were small, but some consistencies appeared. There was a slight tendency for distortion measures to correlate more strongly with movement thresholds than with simultaneity thresholds, and some distortion measures correlated in the same direction with both thresholds. Taken together, these findings suggest that scores for range of apparent movement reflect, in addition to a willingness to perceive the illusion of movement, a possible assimilation phenomenon in which subjects vary in their tendency to stick with a particular judgment (either alternation or movement) once they have made it.

(f) Scanning and Focusing (with Harold Schiffman, Duke University, and Diran Dermen). The dimension of scanning-focusing refers to individual differences in the extent of spontaneous attention deployment. Extensive scanners tend to search and peruse a field of interest, surveying both relevant and irrelevant or miscellaneous properties. The present study attempted to assess scanning and focusing consistencies in a wide variety of tasks in an effort to obtain some convergence of correlations to clarify the interpretation of the construct. These procedures included perceptual judgment tasks (such as size and distance estimation), perceptual speed tasks, and perceptual search tasks (in which the subject must locate in a larger field stimuli of a particular class). Since scanning propensities may also be reflected in the manner in which internal memory fields are surveyed, measures of fluency of ideational production and remoteness of response in word association were also included. A specific attempt was made to differentiate between two possible types of scanning--scanning for signal detection and scanning for information seeking. This was done using perceptual search tasks in which the subject was required to locate stimuli (signals) embedded in meaningfully organized visual fields, e.g., locating faces camouflaged in pictorial scenes. Upon completion of the search task, the stimulus materials were removed, and the subject was then asked specific questions about the content of the scenes. Subjects who incidentally take in information about the field in the process of scanning could thus be differentiated from those whose concern is apparently limited to detecting the signal. Measures were also included for facility in detecting signal stimuli in unorganized fields, such as locating four-letter words in arrays of letters. In view of a theoretical and empirical link between scanning and isolation of affect, several tasks yielding measures of affect expression and control were also included, such as early childhood memories, the Holtzman Inkblot Technique (HIT), and a picture preferences procedure that assesses consistent tendencies to like, dislike, or be indifferent to a wide variety of photographs. These procedures were administered to about 100 male and 120 female college students, along with measures of potentially relevant cognitive styles (particularly field independence and category width), intellectual abilities, and personality.

Factor analyses were performed separately for males and females. Although rotations are not yet complete, some striking relations are discernible.

In males, for example, a factor reflecting incidental knowledge of the content of the field was found to be relatively independent of a factor reflecting facility in locating signal stimuli in unorganized arrays; skill in finding faces hidden in organized fields was associated with both dimensions. In females, a factor emerged linking size estimation with form appropriateness and the absence of shading on the HIT, the latter being a traditional Rorschach indicator of conflict about the acceptance of affect. In both analyses, several cognitive dimensions emerged that displayed substantial correlates with inkblot measures and personality scales.

Another feature of this study was an analysis of errors or "wrong" scores, which was undertaken to investigate the possibility that scanners might be differentiated from nonscanners by the type and number of errors they make. Factor analyses of these "wrong" scores revealed several processes that apparently lead to errors on these tasks. These factors include haste and carelessness, intellectual deficiency, premature perceptual closure, impulsiveness, and susceptibility to distracting and embedding contexts.

(g) Cognitive Styles, Defenses, and Eye Movements (with Lester Luborsky, University of Pennsylvania Medical School, Harold Schiffman, Duke University, and Diran Dermen; partly supported by an NIMH grant to Dr. Luborsky).

Extreme scanning has been found by others to be marginally related to ratings of isolation, projection, and generalized delay on the Rorschach (with isolation possibly related to an "information seeking" type of scanning involved in an obsessive concern with exactness to offset doubt and uncertainty, and with projection possibly related to a "signal detection" type of scanning involved in a paranoid concern with accuracy to offset suspicion and distrust). On the other hand, minimal scanning or focusing has likewise been considered to reflect a concern for accuracy and a preference for viewing the world in a narrow discriminating way. The present study attempts to clarify some of these issues by relating measures of scanning and other cognitive styles (leveling-sharpening and breadth of categorizing) to projective measures of defenses and to characteristic patterns of eye movements in viewing affective and nonaffective stimuli. Heart rate measures of arousal were also obtained under basal control conditions and while viewing the affective and nonaffective pictures.

These procedures have been administered to a small sample of about 40 college students. Several eye-movement scores (such as number of fixations, mean time per fixation, mean track length, and scatter) were obtained for the viewing of each of ten photographs. Heart rate measures were also obtained separately for each photograph presentation. The internal consistency reliabilities of these scores over all ten photographs and over subcategories of photographs reflecting sexual, aggressive, and neutral content have been assessed. Many of the scores proved to be substantially or highly consistent across all of the pictures, but a few were reliable for only one or two of the three subcategories. Such differential reliability was to be expected, since differences in stimulus pattern and meaning should have had more of an impact on some of these scores than on others. The interrelations among eye movement, heart rate, defense, cognitive style, and personality measures have been appraised and a report is in preparation.

(h) Style and Persuasibility (with Fred L. Damarin). This study investigates the interrelations among response styles, cognitive styles relevant to susceptibility to field forces (such as field independence and flexible control), and susceptibility to persuasion and opinion change. A battery of such measures was administered to approximately 90 male college students.

A factor analysis of attitude shift scores, response styles, and selected cognitive and personality measures produced ten factors. Although rotation to oblique simple structure has not yet been completed, preliminary results reveal the presence of four factors clearly related to persuasibility, two of them relatively specific to particular attitude sources and two of them more general in that consistencies are displayed in response to several types of attitude source.

(i) Sex Differences in Cognitive and Personality Structure (with Diran Dermen). This study was designed to bear upon several interrelated questions about the nature and generality of cognitive styles. Of foremost concern were (a) the determination of the number and nature of primary dimensions necessary to account for previously observed stylistic consistencies in cognition; (b) an appraisal of the interrelations of these stylistic dimensions with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; and (c) an evaluation of sex differences in the psychological organization of cognition and personality.

A battery of experimental measures that required over 20 hours of testing time was administered to more than 700 high school juniors and seniors over a two-month period. The sample included approximately equal numbers of males and females, as well as representatives of college preparatory, commercial, vocational-industrial, and art curricula. High school students were sought for this study because of the extreme difficulty of finding comparable samples of males and females at the college level, i.e., samples where either explicit- or self-selection had not occurred differentially by sex on variables relevant to the sex comparison, not only with respect to obviously relevant variables (like ability patterns and interests) but with respect to more subtle ones as well (like maternal education level). The procedures have been scored not only for the specified dimensions under study, but also for various other response properties (such as number of incorrect or unacceptable answers, number of items omitted, and specific response sets) that would help to characterize the test-taking behavior more completely and might provide important controls for clarifying the basis of some of the obtained relations. Since many of the techniques entailed the evaluation of free responses, the scoring itself was a major undertaking that took more than a year to complete. Reliabilities have been assessed on the total sample and separately for male and female subsamples. They have, with few exceptions, proved satisfactory. Correlations have been computed for the same samples. Further analyses are in progress.

(j) The Stability of Cognitive Styles and the Correlates of Change (with Diran Dermen). This study attempts to assess the stability of cognitive style measures over a one-year period and to determine the cognitive and personality correlates of any reliable changes. A sample of 135 high school seniors, who had been extensively tested a year earlier as part of the previous study, were subsequently retested on selected measures of field independence, element articulation, form articulation, conceptual differentiation and compartmentalization, analytic vs. relational categorizing, susceptibility to cognitive interference, and scanning.

Test-retest reliabilities over the one-year interval have been determined. They ranged from zero to .80 with the median being around .50. Differences in corresponding means and variances between the two testing sessions will be evaluated. Base-free measures of change will be calculated for each test, in order to partition the difference in true scores between the two testing sessions into two components: one completely dependent upon initial position in the first session and one reflecting that part of the second performance that is completely independent of the previous performance. The reliabilities of these base-free change scores will be evaluated prior to computing their correlations with each other and with the cognitive and personality measures of the first battery.

In addition, a secondary focus of this study dealt with the malleability of categorizing consistencies. Subjects were required to sort the objects in the object-sorting measure of conceptual differentiation first under standard instructions and then in a new and different way. Facility in changing a natural category system could thereby be assessed, as well as the manner in which the change was brought about.

These indices of flexibility and style in changing conceptual categories will be correlated with cognitive style measures from the concurrent battery and with cognitive and personality measures from the previous battery.

Retest Changes on Personality Scales L. Stricker
(USPHS, 767-2)

PURPOSE: A ubiquitous tendency exists for scores on self-report personality scales to change over time, and in the direction of improved "adjustment," even though no experience intervenes that would be expected to produce such changes. This study seeks to clarify the meaning of these score changes by investigating the nature of individual differences in them. The study will determine the generality of individual differences in score changes on diverse personality scales, as well as the relationships of these individual differences to score changes and initial scores on cognitive and personality measures (not of the self-report type) selected for their relevance in testing hypotheses about the causes of these trends.

PROGRESS: Needed cognitive and personality measures were adapted or developed and administered with the Sixteen Personality Factors Test, the California Psychological Inventory, and the Personality Research Form in a pilot study. The data were analyzed and measures selected for use in the main study. The data for the main study have been collected and the statistical analysis is under way.

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (ARPA, 769-1)

N. Kogan
M. Hermann

PURPOSE: The study has two major aims. The first is to examine the effect of being a representative of a group on behavior in the Prisoner's Dilemma (PD). To date in most research on the PD, subjects have acted in their own behalf, the interaction with the opponent being of major concern. What happens, however, if the opponents in the PD must not only be concerned about each other's behavior but also about how well they are representing a reference group? The second aim of the study is to explore the effects of decision style on responses in the PD. By decision style is meant a subject's way of approaching a decision-making task--e.g., his flexibility, his risk-taking dispositions.

PROGRESS: A preliminary report of the study was presented at the 1969 EPA meetings. A final report has been issued (RB-72-45), and a version for publication is currently in preparation.

MAJOR FINDINGS: The results suggest that whether or not a person is acting in his own behalf or representing a group in the PD will affect how he behaves. Representatives as opposed to individuals tend to be more cooperative to start with in the PD, particularly the highly committed representative. A relationship was established between personality and behavior in the PD. This success may stem from using the dyad instead of the individual as the unit of analysis and from examining the effects of personality across trials.

Study of Aptitude-Achievement Relationships (770-1) J. Carroll

In connection with participation in a conference on the aptitude-achievement distinction that was organized by Dr. Henry Kaiser (University of California, Berkeley) and Dr. Donald R. Green (CTB/McGraw-Hill), Carmel, California, January 1973, Mr. Carroll developed a mathematical model of aptitude-achievement relationships as they apply over multiple school grade levels. This model was an extension of his "model of school learning" as published originally in Teachers College Record, 1963, 64, 723-733. The model postulates that individual differences in aptitude are reflected in different rates of learning, and that these in turn are reflected in measurement achievement levels at different school grade levels. However, measured aptitude is regarded as at least partly or potentially a function of time because of the increased maturity, experience, and "test wiseness" of the examinee. The model makes predictions concerning regressions of achievement on aptitude at different school grades or points of time in the school career of the student. It also makes predictions concerning the means and variances of aptitude and achievement measures. The model was applied to data provided by Dr. George Burket of CTB/McGraw-Hill concerning batteries of aptitude and achievement tests that are vertically equated for grades 1-12 and that were administered to about 175,000 pupils in grades 1-12 on a nationwide scale. The relatively good fit of the model to the obtained data tends to confirm the validity of the model.

Two papers by Mr. Carroll resulting from this work were published as chapters of a book edited by D. R. Green of CTB/McGraw-Hill, The aptitude-achievement distinction, (Monterey, California: CTB/McGraw-Hill, 1974). The chapters are as follows:

"Fitting a model of school learning to aptitude and achievement data over grade levels," pp. 53-78 (originally, RB-73-51).

"The aptitude-achievement distinction: The case of foreign language aptitude and proficiency," pp. 286-303.

Analysis of Data on the Cooperative Literature Tests (770-32) P. Diederich

PURPOSE: The purpose of this project was to identify types of skills, understandings, and appreciations revealed by 19 tests (of two forms each) on the literary works that are most widely taught in grades 9-12. The original intention was to focus on difficulties revealed by these tests, but equal attention is now given to types of items on which students do well and to those on which they do badly.

PROGRESS: The tryout and norming of these tests, in which 211 schools participated in 42 states, yielded nearly 5,000 item-cards with statistical information. A detailed taxonomy of competencies tested by these items has been prepared, and all items (including those discarded after tryout) have been classified under these headings. Examples of items of high and low difficulty in each main category in the tests on each work are being selected. Since the items were submitted by 19 outstanding teachers of these works and reviewed by 38 competent critics, the taxonomy and the sample items should provide a map of what students need to learn, and have learned, about literature in grades 9-12.

Book on Measurement of English Language Arts (770-32) P. Diederich

PURPOSE: The original purpose of this project was to provide the basic text for a correspondence course on measurement of English language arts for teachers of English as a second language in foreign countries. The text circulated for review, however, proved to be more helpful and interesting to American teachers of English than to teachers in other countries.

PROGRESS: All references to the special problems of testing proficiency in English as a second language have been relegated to a supplement that will be included only when the text is used abroad. The American version will be published by the National Council of Teachers of English.

Studies in the Psychology of Language (770-33) J. Carroll

PURPOSE: This project encompasses a variety of approaches to specific problems in the psychology of language, mainly concerned with factors in language production and the analysis of language productions.

PROGRESS: (1) During the previous year's work, a preliminary model of a "performance grammar" for a subset of English sentences was worked out. This model was motivated by the belief that before comprehension processes can be properly studied, it is necessary to develop a theory of how sentences (and utterances in general) are produced from semantic elements assumed to be cognitively prior to their actual encoding into surface structures. It was felt also that rules of the type proposed in transformational generative grammar would be inadequate to account for sentence generation as a behavioral event or for the acquisition of language in the young child. Rather, rules interpretable as behavioral tendencies or learned habits would be necessary. The present model contains a series of contingency rules whereby various kinds of sentences (active, passive, declarative, interrogative, "querying," positive, "challenging," and negative) involving two noun phrases related to a transitive verb may be produced in a left-to-right manner from given semantic elements and variables. The model has been translated into an operating computer program.

While no further progress has been made in expanding this model or in testing it through psycholinguistic experimentation, its possible implications for second language teaching have been described in a paper, "Some suggestions from a psycholinguist," published in the TESOL Quarterly, 1973, 7, 355-367. A more complete description of the model is in press in the International Review of Applied Linguistics.

(2) Mr. Carroll prepared a paper, "Promoting Language Skills: The Role of Instruction," for the 10th Annual Carnegie-Mellon Symposium at Vail, Colorado, June 3-7, 1974. In line with the theme of the symposium, "Cognition and Instruction," this paper made an analysis of the extent to which native and foreign language skills are modifiable by instruction, reviewing relevant literature and offering an interpretation of the possible contribution of an approach through cognitive psychological theory in contrast to two other types of learning theories, naive learning theory and behavior theory. This paper will eventually be published in a book summarizing the proceedings of the conference, under the editorship of David Klahr, Carnegie-Mellon University.

(3) As a member of the Committee on Literacy appointed by the National Academy of Education in response to the late James E. Allen, Jr.'s request to make a report on research and practice in the field of reading, Mr. Carroll has continued his work on editing the final report of the Committee. This report, which will include a number of position papers contributed by leading experts in the field, will be published as a book, Toward a literate society, under the editorship of John B. Carroll and Jeanne S. Chall by McGraw-Hill Inc., in 1975.

(4) A paper prepared several years ago appeared in the 73rd Yearbook, Part I, of the National Society for the Study of Education: "The potentials and limitations of print as a medium of instruction." This paper was an analysis of the extent to which it might be expected that printed prose materials could serve as media of instruction in contrast to other media such as television, direct experience, etc.

(5) Work has continued on the confirmation of the finding, reported previously, (RB-72-10), that age-of-acquisition of words is a more potent predictor of picture-naming latencies than word frequency. The material reported in RB-72-58 appeared in print: J. B. Carroll and M. N. White, "Age-of-acquisition norms for 220 picturable nouns," Journal of Verbal Learning and Verbal Behavior, 1973, 12, 563-576. In the meantime, data collection was completed on an experiment to see whether latencies of picture naming are affected if the subject has prior knowledge of either the word-frequency or the age-of-acquisition classification of the word whose picture he will be naming. These data are being analyzed in preparation for the writing of a report.

(6) Further analyses have been completed on data collected over the period 1971-1973 on absolute pitch judgments -- judgments which can be regarded as a form of retrieval from long-term memory analogous to picture-naming responses. It was established that with respect to accuracy of judgment, persons claiming absolute pitch ability are differentiated from those not claiming such ability in "chroma" judgments (notes within an octave scale) but not in "octave" judgments (judging which octave a stimulus note is in). This finding suggests that pitch judgment is two-dimensional; transmission of octave information carries the approximately 2.5 bits of information found in the case of other sensory modalities, but transmission of chroma information by persons with absolute pitch ability transmits an unusually large amount of information, i.e., up to about 3.6 bits. Latencies of absolute pitch judgments, however, are a function of the total amount of information transmitted (i.e., both octave and chroma). A final report of this research is in preparation.

(7) Further work has been undertaken on developing a mathematical rationale for the asymptotic lognormal word frequency distribution as reported in preliminary form in RB-69-90. With the assistance of Mr. Joseph Deken, a graduate student from Stanford University who spent the summer of 1973 at ETS, the model has been put into more precise mathematical form, and preparations have been made to perform extensive Monte Carlo simulations of the model in order to fit actual data from word frequency distributions.

(8) An article based on data collected some years earlier has appeared in print: J. B. Carroll and J. T. Lamendella, "Subjective estimates of consonant phoneme frequencies," Language and Speech, 1974, 17, 47-59. This report demonstrates that subjects can make fairly accurate estimates of phoneme frequencies, but that these estimates tend to be influenced by perceptions of letter frequencies.

(9) The study performed by Professor Donald Spearritt of the University of Sydney during his tenure as Visiting Research Psychologist at ETS from September 1971 to January 1972, and reported in Research Bulletin RB-72-4,

was published: "Identification of subskills of reading comprehension by maximum likelihood factor analysis," Reading Research Quarterly, 1972, 8, 92-111.

(10) A review of the literature on learning from prose is being prepared as a chapter for volume 3 of the Review of Research in Education. In this connection, it may be noted that Mr. Carroll was Associate Editor for volume 2 of this publication (1974).

Psycholinguistic Theory and Models (770-34) R. Freedle

PURPOSE: The purpose of this project is to investigate language from its earliest inception through adulthood so as to isolate important psychosocial factors which contribute to and modify language development and language use. Particular areas of investigation will include the following topics: (1) the analysis of sequential vocalization behavior of mother and infant as a function of social and situational setting from age 3 months to 2 years; (2) substantive models which will account for the regularities observed in language development from its prelinguistic phase (3 months) to its more formal stage (2 years); (3) the relationships of psychosocial factors on language comprehension, production, imitation, and communication; (4) information-processing models to account for the psychosocial factors on language use; and (5) story comprehension and recall as a function of psychosocial factors as well as a semantic model to describe the internal structure of stories presented for recall.

PROGRESS: In collaboration with M. Lewis the following reports are available: (1) for prelinguistic vocalization in interaction with his mother see RB-72-22; (2) for models of sequential vocalization see RB-71-34; (3) for psychosocial studies see RB-73-18, with William S. Hall; (4) for information-processing approaches to language usage as a function of various psychosocial factors see Freedle and Hall's "An information-processing approach to some problems in developmental sociolinguistics," presented at the biennial meetings of the Society for the Study of Behavioral Development, Ann Arbor, Michigan, 1973; also see Hall and Freedle's "A sociolinguistic account of the influence of peers on the language of black and white children," also presented at the Ann Arbor meetings, 1973. For other information-processing models see RB-73-9 and RB-73-19; (5) for studies of prose comprehension and topic identification as well as message production see Freedle's "Language users as fallible information-processors: Implications for measuring and modeling comprehension" in Freedle and Carroll's Language Comprehension and the Acquisition of Knowledge. Washington, D. C.: V. H. Winston & Sons, 1972.

MAJOR FINDINGS: A Markov model describing six states of the vocalization interactions between mother and infant appears to be moderately accurate in reproducing the sequential states over a two-hour observation period. Further, measures of the transitional probabilities from this analysis appear to be correlated with linguistic measures obtained on the same children at age two years. A full report of the latter findings is in progress. Situational settings significantly affect the sequential flow of vocalizations between mother and infant; furthermore, at different SES levels it appears that the infants may be "learning" different vocalization patterns in interacting with their mothers. Also significant sex differences at this early age can already be noted in vocalization behavior.

An information-processing model for describing the responses made to sentences presented in either standard or so-called nonstandard English indicates that blacks and whites perform equally well when rate of improvement measures for standard English (from age 8 to 10) are employed. In addition, memory storage and retrieval parameters of the information-processing model indicate that the two races are equal in their response to standard English; they differ in which is called a "dialect-preference" parameter which measures the degree to which each population chooses to code the stimulus sentences in either standard or nonstandard form--the latter parameter is characterized as a nonability factor and is seen instead to reflect a number of conflicting cultural pressures to use one dialect over another in a variety of communication settings.

Other information-processing approaches to sentence recall indicated that a serial processing mechanism was favored over a parallel processing one; steps in the information-processing decision tree implicated syntactic and semantic strategies in sentence comprehension for children of mean age 4 1/2 years.

Creativity, Learning, and Attention (771-9) W. Ward

PURPOSE: To develop and validate a nonverbal creativity test suitable for use with preschool children.

PROGRESS: A task has been developed in which children are given a supply of variously colored forms and asked to construct a familiar object, for example, a tree. The task has been included in the test batteries of several large studies of young children which are in progress at ETS. One version, "Make-a-Tree," will be part of the CIRCUS test kit.

MAJOR FINDINGS: The task engages the attention of young children for prolonged periods of time; sometimes evokes highly original constructions; and can be scored reliably and inexpensively for appropriateness, unusualness, and flexibility. Analyses of data collected in the ETS-Head Start Longitudinal Study will provide information relevant to construct validation of these scores.

Development of Cognition and Memory (771-9) W. Ward

PURPOSE: This series of studies is directed toward an understanding of developmental changes in encoding and representational processes. It is proposed that simplistic notions such as Bruner's, that preoperational children can store information only ikonically, are wrong; rather, young children may simply be less able or less skilled at adapting their encoding strategy to task demands.

PROGRESS: Access to information through recognition memory appears to be unaffected by organizational and retrieval strategies which are crucial in tasks involving recall. Recognition tasks, therefore, can provide unconfounded information on encoding and representational processes.

Several studies have been completed in which subjects were shown a large number of pictures to be remembered, followed by a series of yes-no or forced choice tests. Pictures were presented by slide projector; subjects responded by button-pressing, providing measures of both latencies and errors. Variations in instructions and in the dimensions along which target and distractor stimuli differ provide the basis for comparisons which support inferences as to what attributes of presented information are encoded.

MAJOR FINDINGS: One study (RB-73-13) found nursery school and college age students to be highly similar in their relative performance when tested for storage of object identity and color information; these results question the generality of claims that young children are more oriented to color information than are adults. Moreover, little effect was found for instructional variations as to the nature of the subject's activities during learning, suggesting that preferred encoding strategies are relatively resistant to alteration in both young and old subjects. Two further studies, investigating the modifiability of encoding strategy in adults, showed that such subjects can be led by more extreme instructional conditions to adopt primarily either a verbal or a visual encoding approach, but that both kinds of information are encoded and are employed in a recognition test.

Other data, involving latencies in forced choice vs. single choice recognition tasks, provide information useful for the development of models of the information processing routines followed by subjects in recognition tasks. These data suggest that subjects in a forced-choice task engage in several comparison operations which are in fact totally redundant.

Several studies also showed that recognition latencies are shorter for chromatic than for achromatic pictures, and that this effect arises at the time of retrieval rather than as a result of differential encoding: Pictures presented in black and white and then tested in color are recognized as fast as those both presented and tested in color, and faster than those presented and tested achromatically (RB-74-10). Naming latencies are also shorter for chromatic than for achromatic pictures.

Individual Difference Parameters in Recognition Memory W. Ward
(771-9)

PURPOSE: On the basis of recognition memory data, Sternberg has proposed a model of recognition memory in which successive stages of processing involve stimulus encoding, memory search, decision as to a match, and response organization and output. The present study sought to discover whether individuals show consistency across occasions and tasks in characteristics of two of these processing stages, those dealing with encoding and with response organization. It further sought to discover how such individual differences would relate to several popular cognitive style measures.

PROGRESS: College students were given a Sternberg-type recognition memory task on one occasion, and on a second day were given either a similar task or a perceptual matching task which was highly analogous to the memory task. They were also given a cognitive style battery including measures of reflection--impulsivity, analytic style, and color-word reading speed and interference. Analysis is in progress.

MAJOR FINDINGS: Individuals show substantial consistency in the efficiency of the stimulus encoding stage of processing, as measured through the effect on response latencies of stimulus degrading across both testing occasions and type of test. Changes in the probability of positive and negative exemplars, used to index characteristics of the response organization stage, show consistent individual differences across occasions for the same type of task, but not across tasks. On the perceptual matching task, stimulus degrading and response probability effects interact, suggesting that subjects adopt different information search strategies depending on the relative frequencies of the several types of events.

A preliminary report was presented at the 1972 meetings of the Eastern Psychological Association. Project terminated with the publication of RB-73-70.

The Effects of Detoxification on Psychopathology for Alcoholics as Measured by the Differential Personality Inventory (771-24)

D. Jackson

PURPOSE: To assess changes in psychopathology at various stages of alcoholic detoxification and in a treatment program following detoxification.

PROGRESS: The Differential Personality Inventory (DPI) was administered to two groups of consecutively admitted male alcoholics to a state hospital, the first a Detoxification Group, and the second, a group already detoxified.

MAJOR FINDINGS: Of the 27 personality scales, there was a significant reduction in 19 for the Detoxification Group. Scales showing mean decrements included those reflecting Insomnia, Headache Proneness, Broodiness, Cynicism, Depression, Disorganization of Thinking, Familial Discord, Feelings of Unreality, Hypochondriasis, Ideas of Persecution, Irritability, Mood Fluctuation, Neurotic Disorganization, Panic Reaction, Perceptual Distortion, Rebelliousness, Self-Depreciation, Somatic Complaints and Infrequency. Scales reflecting personality and character disorders, such as Hostility, Shallow Affect, and Socially Deviant Attitudes, did not change significantly during detoxification. Other results indicate additional reduction in psychopathology following detoxification, again largely in neurotic and psychotic symptom areas. Data from a DPI validity scale indicated that even patients suffering from acute alcoholic toxicity could answer purposefully. Test-retest data reveal that although systematic mean changes in both groups occur, individuals show a marked tendency to retain their relative position within a group following detoxification and treatment. An article based on these findings was completed in collaboration with Helmut Hoffmann and Paul C. Nelson of Minnesota State Hospital, Willmar, Minnesota, and published in the *Journal of Clinical Psychology*.

Applications of Psychological Scaling Methods (772-11) H. Gulliksen

This paper reviews material on the theory and applications of linear and multidimensional scaling methods. These are methods particularly useful for measuring attributes that have no clear measurable physical correlates-abstract qualities, such as preferences, prestige, beauty, value, which are often difficult to judge. The linear and multidimensional methods, both in the original group average, scale values procedure, and in the individual differences procedure, furnish valuable tools for analyzing and clarifying many problems in psychology, education, and sociology. (A revision of a paper presented at the University of Iowa, Special Seminars in Education Series, February 1973. Copies are available from the author.)

Characteristic Roots and Vectors Indicating Agreement of Data with Different Scaling Laws (772-11)

H. Gulliksen

Given two matrices $X_{n,k}$ and $Y_{n,k}$ of ranks r_x and r_y respectively, (let $r = \min.(r_x, r_y)$ and $0 < r \leq k \leq n$) ;

(a) the nonzero characteristic roots (R) of XY^t and Y^tX are the same;

(b) the right characteristic vectors V of XY^t and W of Y^tX are related by

$V = XWD_x^{-1}$ and $W = Y^tVD_{yt}^{-1}$,

where D_x^{-1} and D_{yt}^{-1} are normalizing diagonal matrices;

(c) the left characteristic vectors L of XY^t and F of Y^tX are related by

$L = YFD_y^{-1}$ and $F = X^tLD_{xt}^{-1}$,

where D_y^{-1} and D_{xt}^{-1} are normalizing diagonal matrices;

(d) $XY^t = V(V^tL)^{-1}RL^t$ (L and V are vectors corresponding to nonzero roots).

Consider four scaling laws; multiplicative, additive (conjoint) measurement, Categorical, and Comparative Judgment.

All four laws necessitate a data matrix of rank two. For categorical judgment only one pair of characteristic vectors is linearly related. For comparative judgment (using the data matrix times its transpose) the two roots are equal and the characteristic vectors linearly related. For the multiplicative and additive cases both pairs of characteristic vectors are linearly related. The multiplicative has two unequal positive roots, the additive has one positive and one negative root.

Lack of fit for any law is indicated by the deviation of the data from the appropriate matrix. (Paper presented at the Psychometric Society meetings in March 1974; available as RB-74-31.)

Comments on "Educational Values for Measurement Technology: Some Theory and Data" (772-11)

H. Gulliksen

Ellis Page and Thomas Breen suggested a procedure for a systematic evaluation of the objectives of education by laymen and by trained educators. The comments stressed the value of comparing this procedure with other more usual methods such as paired comparisons, successive intervals, and multiple rank orders. The desirability of a factor analysis to detect possible different points of view was also emphasized. (Published in William E. Coffman (Ed.), Frontiers of Educational Measurement and Information Systems -- 1973. Proceedings of an invitational conference on the occasion of the dedication of the Lindquist Center for Measurement, University of Iowa, April 6-7, 1973.)

Learning (772-11; NICHHD, 786-9) H. Gulliksen

PURPOSE: A study of learning curves in cats was conducted in an attempt to replicate learning curves in divided brain animals.

PROGRESS: Learning data on normal and split-brain cats has been analyzed by a curve fitting program, developed by James Ramsay, which maximizes the likelihood function for the Audley-Jonckheere learning model by the method of conjugate gradients. A learning model earlier proposed by Gulliksen is a special case of the Audley-Jonckheere model. In the current analysis, values were obtained for the parameters of the model: Rho_0 , Alpha, Beta, Gamma 1 and Gamma 2. The parameter Rho_0 represents the initial probability of a correct response, Alpha represents the amount by which a correct response increments the correct habit strength, and Gamma 1 minus Alpha represents the amount by which a correct response increments the incorrect habit strength. Similarly, Beta represents the amount by which an incorrect response increments the correct habit strength, and Gamma 2 minus Beta represents the amount by which an incorrect response increments the incorrect habit strength.

MAJOR FINDINGS: The values of Rho_0 varied from .28 to .62. A plot of the "left brain" vs. the "right brain" values of Rho_0 did not show agreement of right with left brain values. The values of Alpha were generally small, and the values of Gamma 1 minus Alpha were zero or very small. The values of Beta were smaller than those for Alpha, and values for Gamma 2 minus Beta were very small or zero. Plots of Beta and Gamma 2 minus Beta did not show agreement of left and right brain values. These parameters varied a great deal.

Data from the computer analyses of the learning problems, following the Audley-Jonckheere model and from the conventional learning measures were analyzed to obtain correlation coefficients and other indices of agreement between values for left and right brain learning. Three indices of agreement between first paw and second paw learning were computed:

$$(1) \qquad r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

$$(2) \qquad R_1 = \frac{2\Sigma xy}{\Sigma x^2 + \Sigma y^2}$$

$$(3) \qquad R_2 = \frac{2\Sigma xy - \frac{1}{2N}(\Sigma X - \Sigma Y)^2}{\Sigma x^2 + \Sigma y^2 + \frac{1}{2N}(\Sigma X - \Sigma Y)^2} ,$$

where

$\underline{r}$ is the usual Pearson correlation coefficient. It equals unity if $Y = a\bar{X} + b$ or if $y = ax$, it equals zero if $\Sigma xy = 0$, and it equals -1.0 if $y = -ax$.

$\underline{R}_1$ depends only on deviation scores. It is unity if $X_i - \bar{X} = Y_i - \bar{Y}$ or $x_i = y_i$ for all i, and equals zero if $\Sigma xy = 0$, and equals -1.0 if $x_i = -y_i$ for all i.

$\underline{R}_2$ is an index which equals 1.0 if $X_i = Y_i$ for all i, and equals zero if the variance of $(x + y)$ equals the variance of $(X - Y)$, and equals -1.0 if $X_i = -Y_i$ for all i.

Since it is theorized that learning measures in the split brain animals would be equal for first and second paw performance on the same problem, the values of R_2 should be large if the measures were equal. Where R_2 is small, even though $\underline{r}$ is large, it indicates that though the correlation may be high, the values of X and Y are far from equal. R_1 expresses the relationships between deviations from the mean.

The values of r, R_1 and R_2 were generally near zero indicating lack of agreement between first and second paw performance.

The revised manuscript was issued as an Office of Naval Research Technical report and as ETS Research Bulletin 73-46. It has been accepted for publication by Physiological Psychology. (Study done in collaboration with Dr. T. Voneida of Case Western Reserve University.)

Looking Back and Looking Ahead in Psychometrics (772-11) H. Gulliksen

A brief survey is given of early developments and recent progress in the areas of testing, scaling, factor analysis and mathematical learning theory. (This article is a revised version of a speech presented at the Psychometric Society meetings, March 1972, and is published in the *American Psychologist*, 1974, *29* (4), 251-261.)

Power Scores Estimated by Item Characteristic Curves F. Lord
(772-21; NSF, 816-1)

PURPOSE: A group of students were improperly timed in the regular administration of the SAT Verbal test. The purpose of this study was to estimate by item characteristic curve theory the score each individual would have obtained if he had been allowed to finish the test.

PROGRESS: Item characteristic curve parameters were estimated and power scores computed.

MAJOR FINDINGS: RB-72-48 discusses the method for estimating power scores, giving supportive empirical results, and applies the procedure to 21 students who were mistimed. Published in Educational and Psychological Measurement, 1973.

A Discussion of Certain Reliability Formulas (772-63) W. Kristof

PURPOSE: The accuracies of four different formulas for reliability estimation are compared under certain distributional assumptions.

PROGRESS: The work was presented at the 81st Annual Convention, 1973, of APA and is contained in the *Proceedings*.

MAJOR FINDINGS: There is a uniformly most accurate reliability estimator. Its general use is recommended.

Estimation of Reliability and True Score Variance from a Split of a Test into Three Arbitrary Parts (772-63)

W. Kristof

PURPOSE: The paper gives a method of estimating the reliability of a test which has been divided into three arbitrary parts. The parts do not have to satisfy any statistical criteria like parallelism or τ-equivalence.

PROGRESS: A report was given at the Psychometric Society Meeting in Stanford, 1974. A more complete version will be published in Psychometrika.

MAJOR FINDINGS: If the parts are homogeneous in content (congeneric), i.e., if their true scores are linearly related and if sample size is large, then the method will give the precise value of the reliability parameter. If the homogeneity condition is violated, then underestimation will typically result. However, the estimate will always be at least as accurate as coefficient α and Guttman's lower bound λ_3 when the same data are used. An application to real data is presented by way of illustration. Seven different splits of the same test are analyzed. The new method yields remarkably stable reliability estimates across splits as predicted by the theory. One deviating value can be accounted for by a certain unsuspected peculiarity of the test composition. Both coefficient α and λ_3 would not have led to the same discovery.

On Accuracy in Reliability Estimation (772-63) W. Kristof

PURPOSE: This study in parametric test theory deals with the statistics of reliability estimation when scores on two parts of a test follow a binormal distribution with equal (case 1) or unequal (case 2) expectations. In each case biased maximum-likelihood estimators of reliability are obtained and converted into unbiased estimators. Expressions for the accuracies of the estimators as functions of the reliability parameter and sample size are sought.

PROGRESS: The work is contained in RB-73-24 and has been accepted for publication in Psychometrika.

MAJOR FINDINGS: Second moments of the estimators are obtained and utilized in calculating mean square errors of estimation as a measure of accuracy. A rank order of four estimators is established. There is a uniformly best estimator. Tables of absolute and relative accuracies are provided for various reliability parameters and sample sizes.

On the Statistics of Reliability Estimation Using Parallel Test Forms (772-63)

W. Kristof

PURPOSE: Inferences about the reliability of a given test require repeated measurements in one form or another on a sample of subjects. Two approaches to data collection are common: (a) one obtains multiple measurements using basically the same test whose reliability is the quantity of interest; (b) one obtains multiple measurements using comparable parts of the test whose reliability is the quantity of interest. In the second case the reliability of the component parts is stepped up to give the reliability of the total test. This procedure is not required in the first case. One might therefore assume that case (b) should lead to a statistical theory more complicated than that based on case (a). However, the opposite is true. Too little work with emphasis on statistics has been presented for case (a). Hence it is intended to close a gap in existing mental test theory by developing the statistical reliability theory for case (a).

PROGRESS: A number of pertinent results have been obtained. A Research Bulletin is in preparation.

MAJOR FINDINGS: Suppose that two test forms with bivariate normal score distribution are available. Then the results include the following: (1) The sampling distribution of the maximum-likelihood estimator $\hat{\rho}$ of reliability parameter ρ ; (2) The expectation of $\hat{\rho}$; (3) Bounds for the bias of $\hat{\rho}$ as a function of ρ and sample size N ; (4) The minimal bias of $\hat{\rho}$ as a function of ρ and N . Obtaining these and a number of similar results requires extensive handling of hypergeometric and other transcendental functions. Methods of advanced calculus are needed.

The Generalization Function in Probability Learning Experiments with Infinitely Many Stimuli and Responses

M. Levine

(772-69)

PURPOSE: To discover a procedure for estimating the generalization function in those probability learning experiments in which both stimuli and responses are points on a continuum like length. To study the dependence of the generalization function upon the distribution of reinforcements. To develop methods suitable for studying individual differences in generalization.

PROGRESS: A procedure has been discovered for computing the generalization function directly from individual learning data. The procedure has been applied and found to be very precise. The common translation invariance assumption which had been supported by less precise and sensitive methods was found to be false in both the author's data and previously published data. However, it was discovered that all generalization functions satisfied a functional equation almost as simple as the translation invariance equation. The new equation permits one to analyze generalization functions into two components: one which is invariate over distributions of reinforcements and one which can be deduced (using Weber's law) from the distribution of reinforcements. Although consistent individual differences were observed, all subjects in all conditions conformed to the new equation. Some of these results are reported in a research bulletin now being reviewed and in a manuscript submitted to the Journal of Mathematical Psychology.

Geometric Interpretations of Magnitude Estimation and Other Direct Psychophysical Measurement Procedures (772-69) M. Levine

PURPOSE: To develop a simple, intuitive but quantitative theory capable of integrating the diverse experimental findings arising in the study of direct psychophysical measurement. To account for the data without the ad hoc and implausible hypotheses that have been used by many earlier theories.

PROGRESS: Using only some elementary projective geometry and some intuitions already in the psychometric and experimental psychology literature it has been possible to give a coherent account of many experimental findings. In particular, by assuming that perceptions can be represented as points in a multidimensional space and that orderings of unidimensional attributes are generated by processes like geometric projection one can derive the invariance of ratios of magnitude estimates with change of modulus, some range effects, the concurrence of certain families of magnitude estimation curves (such as those observed in the study of the size-weight illusion) and the regression effect. The theory also predicts a relationship between magnitude estimates and rating scales which seems to be present in published data.

MAJOR FINDINGS: A small number of classical psychological ideas can give a coherent, intuitive and precise account of the major quantitative findings of direct measurement procedures. Details can be found in RB-73-32 which is to appear as a chapter in Contemporary Developments in Mathematical Psychology edited by R. C. Atkinson, D. H. Krantz, R. D. Luce, and P. Suppes. The results on rating scales is being prepared for publication in Psychological Review.

Nonadditive Analogues of the Basic Mathematics of Additive Measurement (772-69)

M. Levine

PURPOSE: The most well-developed area of measurement theory is limited to applications like the measurement of length in which there are no important interactions and in which it is reasonable to assume that the measured objects combine additively. Many of the most important results in this area (additive measurement) depend on a remarkably small number of mathematical results. A type of nonadditive measurement (called affine measurement) is becoming important in modern experimental psychology and psychometrics. The main purpose of this study was to contribute to the development of affine measurement by finding nonadditive analogues of the basic mathematics of additive measurement.

PROGRESS AND MAJOR FINDINGS: Analogues of the key theorems of additive measurement have been discovered. In addition, a simple condition for experimentally testing the appropriateness of affine measurement and a procedure for reformulating nonadditive computations as additive measurement computations have been found. Details are given in RB-73-34 and in the Journal of Mathematical Psychology, 1974, in press.

Parameterizing Curves Transformable to the Same Shape
(772-69)

M. Levine
D. Saxe

PURPOSE: Suppose a family of curves such as the item characteristic curves of a latent trait theory of mental test scores or the isosensitivity curves of signal detection theory can be transformed into curves with the same shape. Then each curve in the family can be identified by a pair of parameters: a mean-like location parameter and a standard deviation-like scale parameter. There are two special cases which are fairly well understood. Case I: The curves are known exactly but the transformation is unknown. In this case, algebraic methods from measurement theory (for example Levine's work on functional equations) can be used to compute the transformation and the parameters exactly. Case II: The curves are known only approximately but the transformation is known exactly. In this case statistical methods (for example Lawton's work on self modeling) can be used to estimate the parameters.

In many potential areas of application, the transformation is unknown and the curves are specified only approximately by data. The purpose of our studies is to develop and test algebraic methods from measurement theory for directly computing parameters from untransformed, empirical curves and to make the methods available in computer programs for general use.

PROGRESS AND MAJOR FINDINGS: Several algorithms have been discovered. They are being tested with Monte Carlo and mental test data. One algorithm involving a novel use of fourier series has been especially successful. When applied to the item test regression curves of the Scholastic Aptitude Test, item difficulties previously computed by Lord using maximum likelihood methods and Birnbaum's logistic model were reconstructed with great accuracy in spite of the fact that no parametric assumptions were made. Some of these results were presented at the Montreal Mathematical Psychology Meetings and the Psychometric Society Meetings. A Research Bulletin describing the algorithm is now being reviewed.

Classification under the Model of Compound Symmetry

(772-79)

L. Gleser

PURPOSE: We are interested in whether the problem of classifying p-variate normal random observations into one of k populations simplifies if it is known in advance that the mean vectors and covariance matrices of the k populations have the structure of compound symmetry first introduced by Votaw.

PROGRESS AND MAJOR FINDINGS: If the mean vectors and covariance matrices of the k p-variate normal populations are known, the form of all admissible classification procedures is obtained. The form of such procedures is identical to the form of the admissible classification procedures in a lower-dimensional subspace (of dimension $r < p$) of p-dimensional space.

When the covariance matrices and mean vectors of the k populations are not known (and have no special structure), but we have sample estimators of these quantities, the classification procedures of Anderson, of Wald, and the likelihood ratio procedures all require distributional results which are extremely complicated. Unfortunately, little simplification results when it is known that the covariance matrices and mean vectors in k populations have the compound symmetry structure. However, in the special case where the covariance matrices and mean vectors have a completely symmetric structure (the intraclass covariance model), enough simplification is obtained so that the needed distributional results are tractable (with the aid of a computer). Expansions for the probabilities of misclassification are given for this special case. The results are being prepared for publication in a Research Bulletin.

Large-Sample Expansions for Null Distributions of Likelihood Criteria (772-79; 772-81)

L. Gleser
I. Olkin

PURPOSE: Many multivariate test statistics (such as the likelihood ratio test statistic for MANOVA) have null distributions whose moments are proportional to ratios of products of gamma functions. For any random variable W, $0 \leq W \leq 1$, whose moments have the above-mentioned form. Box proposed an asymptotic expansion for the cumulative distribution function of W, which provides an accurate method for determining the critical constants defining rejection regions for the multivariate tests mentioned above. Although the method is useful, the calculations needed to obtain the coefficients in each asymptotic expansion must be done ab initio, and almost always involve cumbersome algebraic manipulations. The purpose of the present study is to (i) simplify the needed computations, and (ii) find formulas for the coefficients of the expansion in special cases.

PROGRESS AND MAJOR FINDINGS: Simplified algorithms are given for calculating the coefficients of the asymptotic expansion in the general case. In a certain special case (which includes the null distributions of the likelihood criteria for MANOVA and for testing the independence among sets of variates), explicit formulas are derived for these coefficients. Finally, if $W_1, W_2, \ldots, W_k$ are k statistically independent random variables each of which has a cumulative distribution function which can be expanded in Box form, then it is shown that the cumulative distribution function of $W = \prod_{i=1}^{k} W_i$ also can be expanded in Box form. Formulas relating the coefficients of the expansion for the cumulative distribution function of W and the coefficients of the expansions of the cumulative distribution functions of $W_1, W_2, \ldots, W_k$ are obtained. Two examples of the use of the above methodology to find critical constants for multivariate tests of hypotheses are also provided. A Research Bulletin (RB-72-34) describes the main results of the study. The paper will soon appear in the Annals of the Institute of Statistical Mathematics.

Multivariate Statistical Inference under Marginal Structure. I (772-79; 772-81)

L. Gleser
I. Olkin

PURPOSE: Suppose that we are utilizing k different psychological tests, which have one subtest T_0 in common, and which are designed to be parallel. The components of the g-th test can be represented as (T_0, T_g), where T_0 is the subtest common to all k tests, and T_g is the subtest peculiar to the g-th test. Of particular concern is the hypothesis that these tests are parallel with respect to the means and/or covariances.

PROGRESS AND MAJOR FINDINGS: A complete hierarchy of hypotheses in which the tests are "parallel" has been developed. For example, $H_{m'vc}$ is the hypothesis that the tests are parallel with respect to the means of the common test only, and with respect to the covariances of both tests. This hypothesis might be tested against H_{vc}, the hypothesis of parallelism with respect to the covariances. These hypotheses may be represented as follows:

$$\begin{tikzcd}
 & H_{m'vc} \arrow[r] \arrow[ddr] & H_{vc} \arrow[dr] & \\
H_{mvc} \arrow[ur] \arrow[drr] & & & H_{vc'} \\
 & & H_{m'vc'} \arrow[ur] &
\end{tikzcd}$$

in which an arrow indicates implication. The prime indicates equality for the common test only.

Maximum likelihood estimators under the various models (and under the assumption of normally distributed test scores) have been obtained, as well as the related likelihood ratio statistics. Approximate distributions of the likelihood ratio statistics are worked out, so that the tests can be applied. An example using SAT data is included. This material has appeared as a Research Bulletin (RB-72-40), and a condensed version appears in the British Journal of Mathematical and Statistical Psychology, 1973, 26, 98-123.

Multivariate Statistical Inference under Marginal Structure. II (772-79; 772-81)

L. Gleser
I. Olkin

PURPOSE: The present paper continues the study of statistically testing whether psychological tests are parallel. However, now a more hierarchical structure is permitted. Suppose three tests (perhaps in subsequent years) are used, and suppose further that each test consists of three subtests. We indicate a test by ($\underline{A}$, $\underline{B}$, $\underline{C}$) to denote the parts. In the model considered, the tests have the following composition.

Test 1: (T_0, U_1, V_1)

Test 2: (T_0, U_1, V_2)

Test 3: (T_0, U_2, V_3)

Thus, we see that subtest T_0 is common to all three tests, subtest U_1 is common to the first two tests, and the third part is peculiar to each individual test. The main concern is to test for parallelism with respect to the means and/or with respect to the covariances of the three tests.

PROGRESS AND MAJOR FINDINGS: A likelihood ratio test is developed to test the hypothesis of interest. The asymptotic null distributions of the likelihood ratio test statistic, and of a modified test statistic useful for improving small-sample properties of the test, are obtained. The resulting theory is extended to the case where 6 psychological tests, each composed of 6 subtests, are compared under a design in which tests i and j, $i \leq j$, have the first i subtests in common, $i,j = 1,2,\ldots,6$. This work has appeared as RB-73-64 and will be published in a chapter in a symposium volume on Statistical Design and Linear Models.

Testing If Two Measuring Procedures Measure the Same Dimension (772-79)

L. Gleser

PURPOSE: Mr. F. Lord has suggested a convenient statistical technique for testing the hypothesis that two sets of measurements differ only because of errors of measurement and because of differing origins and units of measurement. Our goal is to obtain the exact level of significance of Lord's test, and to study its power function under various alternatives.

PROGRESS AND MAJOR FINDINGS: Lord has indicated that his test is conservative in the sense that the null hypothesis will be rejected incorrectly less often than the reported level of significance value p of the test would indicate. We have shown that in one sense Lord's test is not really conservative since there is a sequence of parameters in H_0 for which the probability of rejection of H_0 under the test tends to p as a limit. However, there is no single parameter value θ_0 in H_0 for which the probability of rejection is exactly p. These assertions are verified by finding lower and upper bounds for the probability of rejection for H_0. Lord has shown that an upper bound for this probability is p, but our upper bound is everywhere less than his bound. Upper and lower bounds are also obtained for the power function of Lord's test, and the exact asymptotic power of this test is obtained (under different definitions of the term "asymptotic"). The above results will appear in a forthcoming Research Bulletin.

Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study (772-82)

D. Rubin
T. Stroud

PURPOSE: To monitor the level of high school grade standards as measured by the regression line of first year university average on final year high school average. Ontario high school graduates wrote province-wide examinations until 1967, but now the schools are on their own. So far the examining standards of high schools have been fairly uniform, and university admissions offices have been successful at using high school grades as the entrance criterion. But will this continue to be the case if the trends toward more flexibility in curriculum bring about wider divergence in examination patterns across schools? To answer this, we are looking at patterns in data from freshman classes at Queen's University.

PROGRESS: Since the least-squares regression line for students from a given school in a given year is unstable due to the smallness of the data set, it is essential to find a satisfactory method of smoothing the parameter estimates for a given school over a sequence of years (and smoothing across schools is necessary as well). Two methods of doing this have been explored and are described under the headings "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors" and "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors." The random-predictor method has been more successful in that it is considerably less expensive.

MAJOR FINDINGS: More recently a third method is being tried which not only smooths the estimates for school years but incorporates a model for forecasting values for the next year beyond the data set. This method is being tried out on Queen's University data from 85 high schools; actual and forecast values will be compared.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors

D. Rubin
T. Stroud

(772-82)

PURPOSE: To provide a usable statistical model for the study described under the heading "Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study." The model is a random-effects two-way MANOVA with interaction with variables X,Y being the final year high school grade average and first year university average, respectively, and with school and matriculation year as variables of classification. This implies regression lines of Y on X with random intercepts varying over the two-way classification, with slopes and residual variances constant. We wish to estimate the parameters for the purpose of seeing how widely schools differ and with a view to forecasting (with standard errors) the regression parameters for a given school for the current year where university grades are not yet available.

PROGRESS and MAJOR FINDINGS: An empirical Bayes technique has been implemented and run on an APL computer program sequence, using data from 8 schools in 1968, 1969 and 1970. The method is relatively inexpensive and seems to yield reasonable results. The technique features a new matrix inversion algorithm which allows the data set to incorporate an unlimited number of schools. This algorithm is described in a prospective research bulletin.

Approximate Confidence Intervals and Tests for the Gamma Distribution: A Monte Carlo Evaluation (772-82)

T. Stroud

PURPOSE: The gamma distribution is second among continuous distributions only to the Gaussian (normal) in importance in the theory of statistics and applied probability. In psychology it has applications in the theory of learning models and in the analysis of interval comparisons arising from multiresponse experiments. Although standard procedures for large-sample confidence intervals and hypothesis tests are applicable to samples from the gamma distribution, descriptions in the statistical literature of such procedures are fairly rare. In the present research, a number of such confidence intervals and tests are formulated, and then evaluated by Monte Carlo.

PROGRESS AND MAJOR FINDINGS: Univariate large-sample statistical procedures seem to work very well for sample sizes of 100 and moderately well for sample sizes of 25. A proposed bivariate confidence region works moderately well for samples of size 400. All results are contained in a Research Bulletin which is being prepared.

Comparing Regressions When Measurement Error Variances Are Known (772-82) T. Stroud

PURPOSE: In a multiple (or multivariate) regression model where the predictors are subject to errors of measurement with a known variance-covariance structure, we wish to test two-sample hypotheses of (i) equality of regressions on true scores and (ii) equality of residual variances (or covariance matrices) after regression on true scores.

PROGRESS AND MAJOR FINDINGS: Formulas for calculating the test statistic for each hypothesis are presented, based on Wald's asymptotic chi-square procedure. The formulas are cumbersome, but fortunately one need not use them in practice because the value of the chi-square for given data can be obtained using the computer program AUTEST (Lord, RB-72-42, revised edition; Stocking & Lord, RM-73-7). In a study comparing learning in schools in Portland, Oregon, the chi-square was calculated both directly and using AUTEST, and the results agreed to four significant digits. The research has been written up as RB-73-35, and has appeared in Psychometrika, 1974, 39, 53-68.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors (772-82) T. Stroud

PURPOSE: To provide a statistical model for the study described under the heading "Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study." The model is

$$y_{tsi} = (\alpha_t + u_s + w_{ts}) + (\beta_t + v_s + z_{ts})x_{tsi} + e_{tsi} \quad ,$$

where t is the matriculation year, s the high school, and i the student. y denotes university first-year grade average and x the high school final year average. α and β are fixed-effect parameters, and u , v , w , z are random effect parameters. e represents individual variation. There are unknown variance and covariance parameters as well. We wish to estimate the parameters for the purpose of seeing how widely schools differ and with a view to forecasting (with standard errors) the regression parameters for a given school for the current year where university grades are not yet available.

PROGRESS AND MAJOR FINDINGS: A computer program has been written and is working. It computes the variance-covariance estimates using MINQUE. This is the first stage of the estimation procedure. Several difficulties have emerged: (i) computation is expensive, (ii) the estimates have disconcertingly high sampling variances, (iii) only a few hundred students can be accommodated in one run so the data set must be split and run separately, thus lowering the efficiency of the procedure. The program also requires as input a "norm," or starting solution, which must be obtained by some other method. The Lindley-Novick (e.g., RB-70-32) procedure can provide this.

A description has been written up as RB-74-11 entitled "Forecasting a Regression Relationship Which Varies over a Large Number of Subpopulations." The model should have applications to areas other than education, e.g., economics. The method has been described in invited talks at Princeton University, Bell Telephone Laboratories and Virginia Polytechnic Institute.

Estimating Variance Components by Combining Unbiased Estimates from Independent Data Sets (772-82)

T. Stroud

PURPOSE: To generalize the practice of taking the positive part of a negative variance-component estimate to the situation where the parameter has been estimated (unbiasedly) independently in n different sources or data sets. We compare the consistent procedure of taking the positive part of the average of the n unbiased estimates with the inconsistent (i.e., asymptotically incorrect) procedure of averaging the n positive parts. The criterion is mean square error divided by the variance of the unbiased estimator, and the unbiased estimator is assumed to be distributed either as a weighted difference of two independent chi-squares or as the large degrees-of-freedom limit of this, i.e., Gaussian.

PROGRESS AND MAJOR FINDINGS: When the number of sources $n = 2, 3,$ or 4 and when the positive chi-square has 20 or more degrees of freedom, the inconsistent procedure is substantially better, provided the coefficient of variation of the unbiased estimator is less than two. An example where this occurs is the one-way components of variance with, e.g., 26 groups per source and 10 observations per group and a between-group variance of at least 1/35 of the within-group variance.

This research was motivated by the situation in "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors" where the splitting of the data into several computer runs requires the combining of estimates of variance components, some of which may be negative. This research was presented at the 1973 Eastern Regional Meeting of the Institute of Mathematical Statistics, and is available as RB-73-47. It will appear in Journal of the American Statistical Association, June 1974.

Use of INDSCAL in Assessing a Cerebral Deterioration Pattern in Psychiatric Patients (772-82) T. Stroud

PURPOSE: To see if the widely-used INDSCAL algorithm (Individual Differences Scaling--Carroll and Chang) sheds any light on dimension or factor patterns in the Wechsler Adult Intelligence Scale (WAIS) with respect to a population of psychiatric hospital patients.

PROGRESS: The INDSCAL computer program was run using diagnostic groups in place of individuals, and measuring dissimilarity between subtests by sums of squared differences of standard scores over all patients in the diagnostic group.

MAJOR FINDINGS: In the three-dimensional solution, the first dimension has been identified as verbal vs. spatial/quantitative, the second as memory-ingenuity, and the third as "hold-don't hold." The labelling of the third dimension refers to the use of WAIS subtest scores to indicate cerebral deterioration, and is based on contrasting those subtests which deteriorate with age with those which do not. The third dimension loadings showed a correlation of .885 with the subtest weights in a deterioration coefficient due to J. S. Lawson which seems to perform better as a deterioration indicator than does Wechsler's Deterioration Index.

The diagnostic groups that scored as most heavily weighted on the "hold-don't hold" dimension were the senile group and alcoholic psychosis. Alcoholism and drug dependency, on the other hand, were among the least heavily weighted, along with personality disorder, mental retardation, epilepsy and neurosis.

Constrained Latent Class Solutions (772-84)

J. Kearns
D. Thayer

PURPOSE: To find an effective algorithm for estimating the parameters of the latent class model. This approach uses initial results from a general solution for the model and attempts to find a final solution by selecting an appropriate orthogonal transformation matrix subject to inequality constraints.

PROGRESS: The problem was conceptualized as one of optimization subject to both equality and inequality constraints. Equality constraints were imposed on the transformation matrix by means of reparameterization. Inequality constraints were imposed by use of the Fiacco-McCormick unconstrained minimization techniques. It appears that a solution may readily be found for any of several optimization criteria. An appropriate criterion is sought.

MAJOR FINDINGS: A measure of the distance of the final solution from the general unconstrained solution of B. Green was chosen as a criterion to be minimized. Several simulated data samples were generated from a single preselected set of parameter values. Solutions based upon the simulated data gave results which were consistent from sample to sample but were distinctly different from the original solution.

Empirical Bayes Point Estimates of True Score (772-84) J. Kearns

PURPOSE: This study investigates a class of procedures which assign an estimate of true score to every value of the observed score. The estimates are based upon empirical data and approximate the optimal Bayes point estimates. They include (a) "asymptotically optimal" estimates which converge to the Bayes point estimates as sample size increases and (b) estimates based upon various smoothing procedures which reduce the estimation error for smaller samples. Parametric assumptions are made for the distribution of observed score conditional upon true score; none are made for the distribution of true score.

PROGRESS AND MAJOR FINDINGS: Bayes and empirical Bayes point estimates were obtained for the case of the compound binomial error model. A paper was presented at the 1974 Psychometric Society meeting and is available as RM-74-11. Procedures for obtaining estimates with either the Poisson or binomial error models are discussed in Psychometrika, 1973, 38, 533-554 (with William Meredith).

Admissible empirical Bayes estimators, which will minimize the risk function for a class of possible true score distributions, may be obtained if the joint distribution, over the class, of true score moments is known. Generally this distribution is unknown and an improper prior distribution must be selected. A procedure for empirically estimating the quantities necessary for constructing such an estimator has been devised.

On the Evaluation of Empirical Bayes Point Estimates (772-84)

J. Kearns

PURPOSE: Empirical Bayes point estimates which are asymptotically optimal display a certain degree of instability unless the sample size is sufficiently large. This study seeks to provide quick and accurate methods for estimating the distribution of the expected squared error loss incurred when the sample size is fixed. Similar methods applicable to estimates obtained with smoothing procedures are sought.

PROGRESS AND MAJOR FINDINGS: The asymptotic distribution of the overall expected squared error loss has been found for several asymptotically optimal estimators. The corresponding large sample approximations compare favorably with distributions estimated from simulated data. The asymptotic theory has also been applied to estimators which are based upon a smoothing procedure. This work is available in a report coauthored with William Meredith (RB-74-23).

Posterior True Score Distributions (772-84) J. Kearns

PURPOSE: To estimate the conditional distribution of true score for a fixed value of the observed score.

PROGRESS: Estimates of the moments of posterior true score distributions have been obtained for tests which conform to either the binomial, Poisson, or compound binomial error model. A procedure for estimating the stability of estimates of higher moments has been devised. Approximate posterior true score distributions may be obtained by fitting a member of the Pearson family of distributions with the first four moments. A Research Bulletin is in preparation.

A Relative Interaction Criterion for Factor Rotation (772-84)

J. Kearns
J. Balloun

PURPOSE: To provide a new criterion for analytic rotation which explicates the concept of simple structure. The rationale for the criterion is based upon a consideration of the way in which the variance of squared factor loadings is partitioned as in an analysis of variance.

PROGRESS AND MAJOR FINDINGS: The relative interaction criterion was applied to the case of orthogonal rotation using a gradient method of maximization. The results were generally similar to those obtained with Kaiser's varimax criterion. For each of three examples studied the final solutions, considered overall, were somewhat closer than the varimax solution to the subjective graphical solutions reported in Harman's Modern Factor Analysis. A report is available (RB-74-21).

Some Bivariate Convergence Theorems (772-84) J. Kearns

PURPOSE: To provide methods for obtaining the limiting distribution of sequences of pairs of certain linear combinations of random variables.

PROGRESS: An extension to the bivariate case of a theorem originally due to Cramér (1946) has been devised. The theorem provides one with the limiting bivariate distribution of two arbitrary linear combinations of any number of random variables. The theorem requires only that the bivariate distribution of two of the variables, one from each linear combination, converge to a known bivariate distribution and that each of the remaining variables converge to a (possibly different) known constant. The theorem has applications in several areas. A Research Bulletin is in preparation.

Developmental Study of Attention within the First Two Years of Life (NSF, 776-3)

M. Lewis
E. Scott

PURPOSE: This study was designed to study response decrement and recovery in the first two years of life as affected by a number of factors. Repeated and novel stimuli were presented to each subject in two modalities--the auditory and the visual. Each subject received a simple series and complex series in both modalities. The type of change produced by the violation stimulus was changed. Thus, we are able to compare response decrement and recovery across age, across modalities, across complexities, or across stimulus violations. In addition, measures of cognitive development (the object permanence scale of Escalona and Corman and the Bayley Mental Development Scale) were obtained.

PROGRESS: Both a cross-sectional and a longitudinal sample of infants were seen at 3, 6, 9, 12, 18, and 24 months of age. Each subject made two visits--on one the subjects received a visual complex and auditory simple series; on the other, an auditory complex and visual simple (the order was counterbalanced). Measures of fixation, cardiac response, and activity were recorded for the visual episodes. Only cardiac and activity data are available for the auditory episodes. Data analysis is not yet complete for all ages.

MAJOR FINDINGS: A detailed cross-modality comparison of the data collected at three months was carried out. Complexity and sex differences were also considered. The data indicate several trends within and across modalities. These are summarized as follows. For visual stimuli (measured by fixation time), boys showed discrimination on the basis of complexity while girls did not. For auditory stimuli (measured by cardiac deceleration), neither boys nor girls showed discrimination although there is a tendency towards discrimination in the girls' data. All of the discriminations are in the direction of more attention to complex than to simple events. Total amount of cardiac deceleration indicates a sex x modality difference. Girls' deceleration to auditory stimuli is greater than the decelerations associated with any other events, including boys' and girls' visual and boys' auditory. Response decrement across the six trials was found to be minimal in all areas of measurement. There appear to be two important trends in the data: (1) greater stimulus discrimination in the visual than the auditory mode, and (2) greater cardiac response for girls to auditory than to visual events. However, difficulties arise in comparing the nature of the stimuli in the two modalities. It may be that the simple and complex auditory stimuli are not as distinct on the complexity continuum as the visual stimuli. Thus the greater discrimination found in the visual mode may be a function of inherent stimulus characteristics rather than modality differences. These considerations point to the problems that arise in comparisons across modalities. The finding that girls show greater deceleration in the auditory mode than in the visual mode does not suffer from such theoretical difficulties. Our conclusion is in accord with earlier work (Kagan & Lewis, 1965; Watson, 1969) which suggests that such stimulus-modality differences do exist, with girls responding more to auditory than to visual events. Thus comparisons across modalities are informative, but complex and difficult to interpret. Much more empirical evidence is clearly needed before conclusive inferences can be drawn. A paper, presenting the findings to date, was presented at the Eastern Psychological Association meetings in New York in April 1971.

Further analyses were concerned with the questions of the pattern of change of response decrement and recovery across all the ages studied. In addition the questions of individual consistency across age and within and across modalities were explored. Response decrement of the fixation measures to the visual stimuli showed an age effect, with older infants showing more response decrement than younger infants. This result confirms our earlier findings which indicated that amount of response decrement was age related. The second result has to do with stimulus complexity. At each age level except six months there is a complexity effect with the more complex stimuli eliciting less response decrement than the simple stimuli. Again this was not unexpected since stimulus complexity has been repeatedly shown to affect attentional distribution. The decrement in HR deceleration also shows a stimulus complexity effect. Except for 24 months the simple stimulus elicits more response decrement than the complex one.

In general, then, age related response decrement has been demonstrated with two different visual stimuli varying at least in complexity. This age related response decrement appears invariant with respect to the nature of the stimulus. That is, stimulus complexity effects were to be found in the amount of response decrement but not in the age relationship. When the longitudinal data are examined for evidence of consistency across age, the fixation data for the visual data are clearest and indicate a consistency in amount of response decrement across the first 18 months of life. The HR data, both for the visual and auditory episodes, in general, fail to indicate any consistency across age. Given the results of the fixation data we can only conclude that HR during an attentional task is not consistent across age. Whether this is due to the large variability usually encountered in such studies, nature of measurement, or too small sample size remains to be determined.

The longitudinal data were also analyzed to discover if response decrement was consistent within modalities across ages, and also to see if a relationship existed between modalities. The data fail to support within visual modality consistency; however, there is support, albeit limited, for within auditory modality consistency. Across modality consistency in response decrement was strongest for the two episodes which elicited the most regular age effects--namely, visual complex-auditory simple series.

The pattern of response recovery across ages differed with the type of violation introduced. The changes were from straight to curved (curvature); color to black and white (color); and a change in the number of lines. There were three developmental patterns: a linear increase with some asymptotic leveling for curvature, an exponential function for color and a curvilinear function for number violation. This same pattern is visible when looking at percentage of subjects showing fixation recovery.

The HR recovery data by specific violation show little of the pattern that we saw for the fixation data. The only consistent pattern is that for the 3 line curvature change, which shows the greatest HR recovery as it did for the fixation recovery. Once again the cardiac response proves to be only marginally related to the fixation data. The auditory recovery data are comprised only of HR deceleration. Overall the most effective violation is the change in tone from C to F. This change produced the most response recovery while the least effective was the octave change.

Over all ages there was consistency in fixation recovery. Elimination of the 3 and 6 month data when little recovery occurred results in an even greater degree of concordance. Similar results were found for heart rate recovery consistency, especially when the 3, 6, and 9 month data were not considered.

For the auditory episodes the heart rate recovery data failed to show any consistency across age even when the earlier ages were not considered. Tests for within and across modality consistency showed simply that response recovery to one type of violation was not necessarily related to recovery to other types of change.

These findings were presented in the following publications:

Lewis, M., & Scott, E. A developmental study of infant attentional distribution within the first two years of life. Paper presented at the XX International Congress of Psychology, Symposium on Learning in Early Infancy, Tokyo, Japan, August 1972.

Scott, E., & Lewis, M. Developmental changes in response recovery. Paper presented at the Society for Psychophysiological Research meetings, Galveston, October 1973.

Scott, E., & Lewis, M. A longitudinal study of attention and cognition in the first two years of life. Paper presented at the Eastern Psychological Association meetings, Philadelphia, March 1974.

Evaluation of Infant Intelligence: Infant Intelligence Scores--True or False? (NSF, 776-3; Spencer, 815-1)

M. Lewis
H. McGurk

PURPOSE: The purpose of this study was to investigate the predictive value of various tests of infant cognitive development, and the relationships among the various tests. The suitability of administering general tests to evaluate specific infant intervention programs was also considered.

PROGRESS AND MAJOR FINDINGS: Twenty infants were tested at 3, 6, 9, 12, 18 and 24 months. The tasks administered included the mental scale of the Bayley Scales of Infant Development and the object permanence scale from Escalona and Corman's Scales of Sensory-Motor Development. At 24 months, language production and comprehension tasks based on items selected from the Peabody Picture Vocabulary Test were also given. Results indicate that there is no reliable relation between measures of infant intelligence during the first 24 months of life. Relationships among the tests seem to exist only when similar types of items appear on the different tests. No evidence was found to support the view that intelligence is a single unitary capacity which unfolds at a steady rate throughout the process of development. The data also cast doubt on the notion that scores on infant intelligence scales can be generalized beyond the particular set of abilities sampled by the items administered at the time of testing. The implications of these conclusions for evaluation of infant intervention programs seem clear. Simply stated, infant intelligence scales are unsuitable instruments for assessing the effects of specific intervention procedures. This is true primarily because infant intelligence is not a general, unitary trait, but is, rather, a composite of skills and abilities that are not necessarily covariant. The following publications are based on this research.

Lewis, M. (Ed.), *Infant intelligence*. New York: Plenum Press, 1974, in press.

Lewis, M. Infant intelligence tests: Their use and misuse. *Human Development*, 1973, *16*, 108-118.

Lewis, M. & McGurk, H. The evaluation of infant intelligence: Infant intelligence scores--true or false? *Science*, 1972, *178*(4066), 1174, 1177.

Developmental Model of Cognitive Style (779-3)

A. H. Baker

Study 1: Relations between cognitive controls and school performance in the early grades.

PURPOSE: To identify principles by which children organize and process incoming information from the environment and to relate these individual differences in perceptual-cognitive style to academic success in the early school years. This research is being framed within the context of a particular developmental model, according to which all cognitive controls are not equally available at all stages of life, and a newly formed cognitive control does not replace one that was formed earlier. Rather, the model views cognitive controls as forming a hierarchy of organizations of cognitive behavior, and makes use of the concept of levels within each organization. Focal Attention (i.e., directing attention actively to stimuli and scanning broad sections of the environment so that unique properties of objects are registered) is assumed to emerge first in development. The principle of Field Articulation (i.e., attention directed selectively to relevant stimuli and withheld from irrelevant stimuli guided by the central requirements of the task) emerges next, subordinating and integrating Focal Attention as a distinct process. Leveling-Sharpening (i.e., maintaining in memory the organization of relevant-irrelevant information in the form of a differentiated image with which subsequent information is related and compared) is the third cognitive control principle to emerge and is assumed to subordinate and integrate the preceding two principles. Finally, Equivalence Range is the fourth principle whereby symbols or labels are introduced in order to interrelate bits of information allowing objects to be dealt with in terms of categories or classes. In brief, the present research is aimed at exploring how each of the above described cognitive style dimensions relates to school performance in kindergarten and first grade children.

PROGRESS: One hundred fifty white, low to middle class kindergarten children have been given a battery of procedures which assess the above-described cognitive controls, prereading and reading skills, ratings of general classroom behavior (by the teachers), impulse control, and selected personality measures.

It is planned to relate a child's cognitive developmental status, in terms of information processing strategies, with the status of his cognitive skills in terms of reading. The goal is to search for patterns of information processing strategies (and their level of development) associated with the emergence of effective reading skill and to learn more about the relation of certain aspects of personality here presumed to be critical in the learning process and the early formation and operation of cognitive styles.

Fifty first grade children who had been evaluated one year ago were also recently given the same battery of procedures outlined above. The focus here was upon (a) the stability or change in patterns of the cognitive structures and skills under study, and (b) the interrelationship between the assessments of academic success obtained from the teacher and the cognitive status observed during the present testing as well as that observed a year ago.

Data collection and processing phases have been completed for both aspects of this study and data analyses are now in progress. (Study done in collaboration with Sebastiano Santostefano, Boston University School of Medicine.)

Study 2: Individual differences in attentional processes in adults and children.

PURPOSE: Recent research has examined the hypothesis that individuals can be ordered along a dimension in terms of their tendencies to reduce or to augment the intensity of persistent incoming stimulation. At one extreme, the reducer attenuates the intensity of incoming stimulation, thus enabling him to handle high intensities of stimulation quite well but making him acutely uncomfortable when environmental stimulation is minimal (e.g., under sensory deprivation). At the other extreme, the augmenter magnifies the intensity of incoming stimulation, thus showing intolerance of high levels of stimulation but high capacity to cope with very low stimulus intensity. In between is the moderate, who neither greatly enlarges nor minimizes incoming stimulus intensity. Measurement of this dimension has involved variants of the Kinesthetic Figural Aftereffects Task (KFA). Individuals are designated as reducers, as moderates, or as augmenters if a standard block is judged respectively as smaller, as about the same size, or as larger after intervening stimulation. Findings have indicated, for example, that reducers are more tolerant of pain, less able to endure sensory deprivation, more likely to show reduced averaged cortical and visual evoked responses to strong light stimulation, and more likely--if male--to participate in "contact" athletics such as football.

In a series of interrelated studies, the present research seeks to explore the following substantive-theoretical and methodological issues suggested by prior KFA research: (1) Most differential studies utilizing KFA procedure can be criticized either on the basis of the psychophysical procedures employed and/or on psychometric grounds. The first phase of the present research will therefore simultaneously reexamine the reliability and validity of two major variants of the KFA procedure, using normal adults as subjects, and if successful, (2) an attempt will be made to delineate more precisely the nature of the psychological processes underlying individual differences in performance on KFA tasks. (3) A third phase will attempt to delineate educational implications of the augmentation-reduction dimension. The augmentation-reduction typology suggests that the optimum learning environment may be quite different for these three kinds of individuals. The reducer should learn best in an environment in which the techniques that introduce educational materials utilize fairly high levels of sensory stimulation, in which a variety of educational approaches are utilized and in which stimulation is provided by frequent social interaction. However, a calm, nonstimulating environment should best promote learning by the augmenter, who needs far less stimulation than does the reducer and, in fact, would be quite uncomfortable in highly stimulating surroundings. Augmenters should also profit from educational projects which they are given to carry out on their own. (4) Finally, this research will focus upon age changes in the hypothesized augmentation-reduction dimension, and in the correlates of this dimension, with special reference to other possible dimensions which may be tapped by the KFA task.

PROGRESS: Since its inception in the summer of 1970, the following has been achieved: (1) an extensive literature search has been accomplished; (2) a number of substantive and methodological issues which appear to be critical in studying individual differences in kinesthetic figural aftereffects have been tentatively identified, and a research program aimed at exploring them has been broadly sketched; (3) equipment has been secured for administering the two principal variants of the KFA task; (4) an extensive battery of tests and tasks aimed at exploring the issues described above has been assembled and/or developed; (5) personnel have been recruited and trained in the administration of the two types of KFA tasks and of the test batteries; (6) ten different studies, aimed at exploring the goals of the present research delineated earlier, have been completed.

The findings to date have both clarified a number of substantive and methodological issues, while at the same time expanding our conceptualization regarding this research area. Substantively, it is now clear that individual differences in the KFA task reflects both trait variance and state variance. Thus, as a personality dimension, KFA scores predict individual differences in thinking about the future, time perspective, and attitudes toward death, findings which hold not only for an undergraduate population but also for a community-active geriatric group. At the same time, however, a number of aspects of organismic state--including the degree of self-rated tiredness, day in the monthly menstrual cycle, intake of various medications, presence or absence of illness or pain etc.--affect scores on the KFA task. Methodologically, it has been shown that when such "state" variables are controlled, there is a significantly greater positive test-retest reliability than when state is not controlled. Moreover, certain problems in the use of a test-retest design in the study of kinesthetic figural aftereffects have been delineated.

In expanding our interests regarding the ways in which people react to the intensity of incoming environmental stimulation, we have extended our attention beyond the Kinesthetic Figural Aftereffects task. As a result, our efforts during the past year have focused on devising a very extensive battery of tasks which would assess this dimension in multiple modalities using multiple methodologies, while at the same time including in this battery a number of the more traditional cognitive style variables to clarify possible interrelationships between the present dimension and the dimensions studied by other researchers. As a result of this effort, we have devised a highly diversified but very specialized 12-hour battery of tests which includes not only extensive paper-and-pencil measures but a wide variety of psychophysical tasks pertinent to the present problem area. This battery has been administered to 120 college-age subjects; all data has been processed, and data analyses are in progress.

Another line of investigation has focused upon individual differences in the aging process. Because the reducer tends to damp down incoming stimulation, it has been hypothesized and found in earlier studies that he shows greater tendency to seek more intense and more complex levels of stimulation. We therefore hypothesized that (1) reducers would show a higher level of social involvement and (2) the reducer would maintain his higher level of social interaction during old age whereas the augmenter would show more tendency to withdraw from social interaction. Both predictions have been confirmed in a recent self-report type study.

Results of some of the above findings have been or will be reported as follows: (1) Individual differences in the psychology of death. To be presented at meetings of the American Psychological Association, August 1973, Montreal, Quebec, Canada. (2) Effects of variation in ongoing and induced organismic state upon kinesthetic figural aftereffects. Presented at meetings of the Eastern Psychological Association, May 1973, Washington, D. C. (with Brian Mishara, Irene W. Kostin, and Laurence Parker). (3) Will to live, future planning, and kinesthetic figural aftereffects: Towards a theory of individual differences in the aged. Presented at the 25th Annual Meeting of the Gerontological Society, December 1972, San Juan, Puerto Rico (with Brian Mishara and Irene W. Kostin). (4) Do people who seek less environmental stimulation avoid thinking about the future and their death? A study of individual differences in kinesthetic figural aftereffects. Presented at meetings of the American Psychological Association, September 1972, Hawaii (with Brian Mishara and Irene W. Kostin). (Study done in collaboration with Dr. Brian Mishara of Socio-Technical Systems Associates, Boston.)

The Integration of Faces and Voices in Infancy
(779-4)

M. Lewis
L. Rosenwein
H. McGurk

PURPOSE: Two studies were done under this rubric. The general purpose involved the question of whether intermodality integration exists in early infancy. Most theories of perceptual development postulate a gradual intermodality integration with age. We were intrigued by Aronson and Rosenbloom's (Science, 1971) research showing that one-month-old babies were disturbed by a mismatch between their mothers' faces and the displacement of their voices. The first study was designed to replicate this finding in a more controlled experiment. The second study was designed to elaborate and extend theoretical issues investigated in the first study. We wanted to know whether learned information can be coordinated between two senses. In order to determine the answer to this question we explored the infant's reaction to a stranger's voice mismatched with his mother's and vice versa as well as to face and voice mismatches involving location.

PROGRESS AND MAJOR FINDINGS: Three publications outlining the work on intersensory integration are available:

Lewis, M., Townes-Rosenwein, L., & McGurk, H. Normal and discrepant face-voice integration in early infancy. Research Bulletin 74-4. Princeton, N. J.: Educational Testing Service, 1974.

McGurk, H. & Lewis, M. Space perception in early infancy: Perception within a common auditory-visual space. *Science*, in press.

Townes-Rosenwein, L. & Lewis, M. Normal and discrepant face-voice integration in early infancy. Paper presented at the Eastern Psychological Association meetings, Philadelphia, April 1974.

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development

M. Lewis
C. Wilson

(779-4)

PURPOSE: This longitudinal study investigates attentional processes during infancy and their relationship to cognitive growth in the preschool period. Specifically, the study explores: (1) problems in measuring attention; (2) factors influencing attention, such as sex, and clinical problems (birth trauma); (3) stimulus differences such as facial vs. nonfacial and familiar vs. novel; (4) developmental consequences of differential attention, including the relationship of attentional differences in infancy to preschool performance in concept formation tasks, learning problems, reading, and language acquisition.

PROGRESS: To explore these problems, two samples of 32 infants, 16 boys and 16 girls, were seen at 6, 13, 25 and 44 months of age. At each age level, a variety of visual and auditory stimuli were presented to the infants and a large number of response measures were taken. In addition to their response to visual and auditory stimulation, a wide variety of other behaviors were observed. These included such behaviors as concept formation, problem solving, task persistence, language acquisition, play behavior, and IQ measures.

The data, yet to be fully analyzed, will be used to explore the major areas of interest outlined below.

(a) Measurement of attention. The responses observed have been: (1) fixation (receptor orientation), (2) vocalizing, (3) crying, (4) autonomic reactivity (specifically, heart rate and respiration) and (5) gross motor activity. Investigators often harbor implicit assumptions about each of these response indices. For example, differential fixation time is often assumed to reflect a capacity for differentiation or discrimination. More significantly, absence of differences in fixation time is believed to reflect the infant's inability to differentiate among the visual patterns being presented. This inference may be fallacious, for often the infant fixates for an equal period upon two stimuli but will only vocalize or smile to one of them. Vocalizing has been assumed to indicate both a specific social response and general arousal level. Similarly, gross motor activity or heart rate acceleration is often regarded as an indicator of activity or arousal level. It is clear that detailed empirical work is necessary to assess the viability of these assumptions. At present, it appears that each of the infant's responses viewed singly is probably ambiguous in meaning and can reflect different psychological processes. A long fixation time can occur to a familiar stimulus or to an uncertain one. Cardiac deceleration can occur when the infant studies the stimulus and is quiet, but acceleration often appears when he studies the stimulus and is active. Thus the present study intends to contribute information relevant to understanding the varied responses associated with distribution of attention to auditory or visual patterns. Analysis of six-month data indicates that attention to visual and auditory stimuli is usually associated with the following type of response pattern: receptor orientation; decrease in gross motor activity, crying, and vocalizing; autonomic responses of cardiac deceleration; and a slowing of respiration. Individual differences in this response pattern have been observed. Developmental differences in this response pattern have yet to be explored.

(b) Individual differences in attention. Marked individual differences have been found in the attentional behavior of six-month-old infants. Are these differences consistent in the visual and auditory modalities, and are the differences stable over age?

(c) Stimulus differences in attention. Because a variety of visual and auditory stimuli were presented at each of the four age levels, it will be possible to explore the influences of different stimuli on attentional behaviors. Both visual and auditory stimuli can be divided into social and nonsocial, i.e., human forms, faces, or voices versus nonhuman patterns or sounds. These two classes of stimuli were presented at all age levels and the study seeks to explore the effects of these stimulus qualities on the attentional process, i.e., their ability to attract (length of time oriented toward) the infant's attention, and the ability to elicit social responses, i.e., vocalization and smiling.

(d) Relationship between attentional measures and cognitive capacities and styles. Along with the various attentional measures to visual and auditory stimuli, a variety of other cognitive capacities and styles were observed for the one-, two-, and three-year levels. These include: (1) language tests at two and three years, (2) play behavior at one, two and three years, (3) concept formation at two and three years, (4) problem solving at two and three years, and (5) IQ scores at three years. The cognitive style measures of impulsivity and reflectivity as measured by MFF task and persistence task were given at three years. It is therefore possible at each age level to investigate the relations among these various measures of cognitive capacity and to relate them to the attentional measures. For example, there is evidence that amount of fixation is related to play behavior at one year and that IQ scores can be predicted from attentional scores at the three-year level.

(e) Personality development. In the longitudinal investigation of attention and cognitive growth, several personality characteristics were assessed. Observation of individual differences in attentional behavior revealed that personality characteristics might covary with them. Personality measures obtained at the one-, two- and three-year levels were: (1) measurement of the infant's vigor, (2) dependence behavior as measured by interaction of infant with mother or teacher, (3) sex role appropriate behavior measured in play behavior, and (4) task persistence. Preliminary observations suggest that a dependent child may spend less time looking at nonsocial stimuli as compared to social stimuli. Moreover, when left alone in a relatively unfamiliar room, his attention span might be more a function of wanting to see the mother than of interest in the stimulus being presented. It was also observed that vigorous children had more difficulty attending than placid, less vigorous ones who showed little physical activity. Analysis of the data is proceeding along all domains of interest.

MAJOR FINDINGS: (a) Measurement of attention. Principal component analyses on the attention data indicate that the configuration of two components, together accounting for more than half the total variance, is quite consistent both for the different types of attentional stimuli and across age. The first component has been called an orienting component of attention, as it represents those responses involved in orienting to the stimulation: fixation, cardiac deceleration, and a reduction in gross activity. The second component is interpreted as an affect component, representing the vocalization and smiling responses to the attentional stimuli.

(b) Individual differences in attention. In general the results show little consistency of individual differences in attention, either between the visual and auditory modalities, or across the four age points. No consistent sex differences were found. It appears that the method of assimilating environmental information varies as a function of the type of information to be assimilated. Moreover, patterns of attention vary with age, perhaps in the search for the most efficient method of processing incoming information.

(c) Stimulus differences in attention. The analysis of the data for the social and nonsocial stimuli separately revealed few consistent differences. However, principal component analyses on the responses to the two kinds of stimuli indicated a stronger (accounting for a greater percent of the variance) orienting component for the nonsocial than for the social stimuli, and a stronger affect component for the social than for the nonsocial stimuli. Thus, the difference in the response to the two types of stimuli is not so much a difference in quantity as in the relative predominance of certain response measures such as vocalizing and smiling.

(d) Relationships between attentional measures and cognitive capacities and styles. There appears to be no consistent relationship between the attentional and cognitive measures, indicating that the processing of incoming information (attention) is separate from cognitive capacity. In view of the lack of consistency in patterns of attention it is not surprising that the attentional and cognitive measures are not related. Even if cognitive capacity is assumed to be constant (a tenuous assumption at best), the variations in attention as a function of both type of stimulus and age preclude the finding of consistent relationships between attention and cognition.

Some of the results have been summarized and appear in the following reports:

Wilson, C. D., & Lewis, M. A developmental study of attention: A multivariate approach. Paper presented at the Eastern Psychological Association meetings, New York, April 1971 (also Research Bulletin 72-31. Princeton, N. J.: Educational Testing Service, 1972.)

Wilson, C. D., & Lewis, M. Temperament: A developmental study in stability and change during the first four years of life. Research Bulletin 74-3. Princeton, N. J.: Educational Testing Service, 1974.

Mothers and Fathers, Girls and Boys: Attachment Behavior in the First Two Years of Life (779-4)

M. Lewis
M. Weinraub
P. Ban

PURPOSE: In this longitudinal study of one- and two-year-old children, the child's expression of four attachment behaviors--touching, looking at, vocalizing to, and proximity to each parent--was examined. We were interested in observing (a) how these four behaviors are interrelated both within and across ages, (b) sex of child and sex of parent differences in the use of attachment behaviors, and (c) the relationship between the patterns of attachment behavior expressed to each parent and measures of the child's intelligence.

PROGRESS AND MAJOR FINDINGS: A small group of children was observed individually at one and two years of age in a naturalistic free play situation in the presence of their mothers and again in the presence of their fathers. Intelligence tests were administered to the children.

There were no overall child sex differences in the expression of attachment behaviors toward the parents at either age. There were parent differences at age one, with more attachment behavior directed toward the mother than toward the father. By two years of age these parent differences in the expression of sex differences disappeared for the most part.

The data suggest that, in general, there may be a transformation from the use of proximal (touching and proximity) attachment behaviors to the use of distal behaviors (looking and vocalization). However, both the sex of the parent and sex of the child influences the clustering of attachment behaviors. Measures of infant intelligence were found to be correlated with the changing patterns of attachment. More intelligent children tended to show greater decreases in their expression of proximal attachment behaviors and greater increases in their expression of distal attachment behaviors than less intelligent children.

The sex of child and sex of parent differences we observed in the expression of attachment behaviors by two-year-old children seem to be precursors of sex differences in interpersonal behaviors characteristic of older children and adults. These findings have been presented in the following publications:

Ban, P., & Lewis, M. Mothers and fathers, girls and boys: Attachment behavior in the one-year-old. Paper presented at the Eastern Psychological Association meetings, New York, April 1971. Merrill-Palmer Quarterly, in press.

Lewis, M., Weinraub, M., & Ban, P. Mothers and fathers, girls and boys: Attachment behavior in the first two years of life. Paper presented at the Society for Research in Child Development meetings, Philadelphia, March 1973.

Lewis, M., & Weinraub, M. Sex of parent x sex of child: Socioemotional development. Paper presented at the International Institute for the Study of Human Reproduction Conference on Sex Differences in Behavior, Tuxedo, New York, October 1973. Also, to appear in R. M. Richart, R. C. Friedman, & R. L. Vande Wiele (Eds.), *Sex differences in behavior*, in press.

Mother-Infant Interaction from the First Day to the Third Month of Life (779-4)

M. Lewis
S. Lee-Painter

PURPOSE: To explore the earliest relationship of the neonate and its mother: the period immediately after birth; to relate this relationship to the later interaction between infant and mother.

PROGRESS: Thirteen neonates and their mothers were observed in a London hospital on the first, fourth, and fifth days after birth. Observation data on 24 neonatal and 24 maternal behaviors were gathered for periods up to one hour. In addition, the Brazelton Behavioral and Neurological Assessment Scale was administered. The same infants were seen in their homes for naturalistic observation of up to one hour at one month, two months, and three months of age. At these observation points, 13 infant and 13 maternal behaviors were observed.

MAJOR FINDINGS: Data analysis is now in progress. The results from this project will be valuable as a cross-cultural comparison to similar data available from previous studies on mother-infant interaction.

Differentiation and Development of Social Norms W. Emmerich
(779-7)

PURPOSE: The study examines developmental trends in age-role and sex-role normative differentiations and relates these trends to developmental theories of social learning, cognitive complexity, and cognitive structure.

PROGRESS: Shared beliefs on how the self should act toward others (norms) were studied cross-sectionally in 680 middle-class children and adolescents. Subjects judged the normative import of statements differing in (1) behavioral content (agree with other, help other, seek help from other, argue with other); (2) sex and generation of reference figures holding norms for the self including the self (sources), and (3) sex and generation of recipients of the self's actions (objects).

MAJOR FINDINGS: The four behavioral contents were differentiated similarly at all ages (continuity), but age-role and sex-role normative differentiations of sources and objects exhibited marked developmental trends (discontinuities), often generalizing across behavioral contents. During middle childhood higher standards were expected in relationships with parents than with peers irrespective of source, whereas during adolescence this differentiation was reversed for attributions to sources within the child's generation. Sex-role norms were strongly sex typed during middle childhood, a pattern that decreased during later childhood and adolescence. Related developmental shifts occurred in choices of identification models, indexed by convergences between norms held by the self and those attributed to other sources. Certain developmental trends were accelerated in brighter subgroups. The study is published in the Journal of Personality and Social Psychology (with K. S. Goldman and R. E. Shore), 1971, 18, 323-353.

The study provided information on developmental trends in extreme response choice, found to decrease linearly with age in both sexes and earlier (curvilinearly) in girls. Cognitive level was found to mediate choice of an absolute and extreme category, but not choice of a non-absolute extreme scale category. This aspect of the study was published in Developmental Psychology, 1971, 5, 540.

In a follow-up study, middle-class children and adolescents evaluated single trait descriptions of hypothetical persons of their own sex. Developmental trends in trait evaluations were examined and related to the subject's sex, cognitive ability, and generational status of the hypothetical stimulus person (adult vs. peer). Certain findings were most consistent with ego-developmental theory, while others were more consistent with a person-perception analysis or with a social desirability interpretation. Popularity and similarity judgments were partially differentiated from other components of value such as like-dislike. Evidence suggested that "personal" evaluations become increasingly differentiated from "social" desirability judgments with development, and that self-esteem is maximally engaged when the individual evaluates himself on traits defining stage-related themes in ego development. This study has been published in Child Development, 1974, 45, 172-183.

Findings from the above studies related to sex-role development are incorporated in a chapter entitled "Socialization and Sex-Role Development" to appear in P. B. Baltes and K. W. Schaie (Eds.), Life-Span Developmental Psychology: Personality and Socialization. New York: Academic Press, 1973.

Center for Child Care Research (779-8)

I. Sigel
R. Cocking
B. Bryant

PURPOSE: To study the application of a cognitive-affective model of preschool education to a group of four-year-old children representing middle and working class backgrounds. Our interest is in determining the more effective method of application, a small group learning situation or a large group, and related interest in method of instruction. We are interested in determining whether we can evolve a curriculum meeting program objective of enhancing representational skills through the use of play.

PROGRESS: Although no children have been selected as yet, we are in the process of developing assessment procedures, curriculum materials, and the physical space. We have also begun to create a master list of children born within close traveling distance to ETS from vital statistics records and will soon be working out arrangements to draw a sample. At present this program is supported by ETS, but efforts are being made for outside support.

Cognitive Styles and Psychopathology (NICHHD, 782-3) S. Messick

PURPOSE: Recent studies of perceptual and cognitive response patterns of schizophrenics have indicated striking differences between groups of patients with different symptom patterns (paranoid vs. nonparanoid), different levels of adequacy in life adjustment prior to illness (good premorbid vs. poor premorbid adjustment), and different lengths of institutionalization (acute vs. chronic). The present study attempts to utilize the extreme forms of cognitive organization and various personality trait patterns.

PROGRESS: An extensive battery of perceptual, cognitive, and personality measures has been administered to a sample of 95 hospitalized schizophrenics. These procedures were administered either individually or in small groups of four or five patients. The battery included measures of perceptual field-independence (Rod-and-Frame Test, Embedded-Figures Test), element articulation, form articulation, scanning, verbal closure, perceptual closure, extremity and confidence of judgment, category width, risk taking, leveling-sharpening, conceptual differentiation (object sorting), ideational fluency, and verbal comprehension. Also included were items of case history information; a personality inventory assessing several dimensions of neuroses and character disorder; and several rating scales for symptom expression, properties of delusions, and suddenness of onset of illness.

The sample contained roughly equal numbers of paranoids and nonparanoids, who in turn were approximately equally divided into acute and chronic subgroups (using 18 months of hospitalization as the dividing point). The scoring of these measures has been completed, and their reliability for this sample has been appraised and found satisfactory. Various analyses are under way to evaluate subgroup differences on cognitive style and personality measures (in a 2 x 2 breakdown of paranoid vs. nonparanoid and acute vs. chronic patients) and to isolate patterns of cognitive style-personality trait interrelations. (Study done in collaboration with J. Silverman of NIMH and D. N. Jackson of University of Western Ontario.)

Longitudinal Study of Creativity (NICHHD, 782-5) N. Kogan

PURPOSE: To examine the predictive validity of the Wallach-Kogan divergent-thinking tasks over a seven- and a two-year period. Subjects in two school systems who had been assessed in the fifth and tenth grades were given the Wallach-Wing questionnaire of nonacademic attainments upon graduation from high school.

PROGRESS: Two earlier reports have been published (Journal of Personality, 1968, 36, 420-439 and Child Development, 1972, 43, 427-442). A final report has also been issued (RB-73-71) and is in press in the Journal of Educational Psychology.

MAJOR FINDINGS: Fifth-grade ideational productivity had no predictive power in respect to nonacademic attainment at graduation, whereas tenth-grade ideational productivity made a marginally significant contribution to that criterion in one of the school systems. Assessment of nonacademic attainments was found to be fairly stable from tenth-grade to graduation. Examination of separate fields of nonacademic attainment indicated diverse relationships with cognitive assessments.

Cognition of Affect (NICHHD, 783-2)

S. Messick

PURPOSE: This study investigates individual differences in the conceptual differentiation of affects as a function of cognitive styles and ideological commitment.

PROGRESS: Several measures of cognition and personality have been administered to over 600 high school students along with a task in which they were asked to judge the degree of similarity or difference among 18 affect names, such as fear, surprise, joy, and contempt. These judgments of similarity were analyzed by multidimensional scaling procedures to isolate consistent individual viewpoints about affect interrelations. The first two dimensions of judged similarity clearly reflect positive vs. negative affect, and intensity (calm vs. excited). The third dimension is tentatively interpreted as reflective of negative feelings about the self (ashamed, humiliated, afraid) vs. negative feelings directed externally (contemptuous, mad, angry). This last dimension may correspond to intropunitiveness vs. extrapunitiveness. The number of stimulus dimensions and other structural properties of each viewpoint will be used to assess consistent individual differences in the conceptual differentiation of affects. The content of the stimulus representation for each viewpoint will be used to characterize the cognition of affect according to that viewpoint. Individual differences in conceptions of affect will be studied in relation to the cognitive and personality measures included in the battery, but in particular to measures of skill in discriminating and identifying visually displayed affects, the spontaneous tendency to characterize photographs of faces in affect terms, choice of preferred affect, and ideological polarity. The scoring of these procedures is currently in progress. (Study done in collaboration with S. S. Tomkins, City University of New York.)

Dimensions and Categories of Physiognomic Judgment N. Kogan
(NICHHD, 784-2)

PURPOSE: The general purpose of the present study is to examine the manner in which the use of an age judgment scale varies with the chronological age of the judge.

PROGRESS: A report (RB-73-68) has been issued and is in press as a Brief Report in Developmental Psychology.

MAJOR FINDINGS: A main effect for judge's age was obtained only in the case of the male photos--older judges attributing higher age levels to the stimulus persons. For both male and female photos, a highly significant interaction (age of judge by age of stimulus person) was obtained. Stimulus persons at the younger end of the age continuum were judged consistently older by the older judges. The inflection point is approximately age 40 for males and age 30 for females. Stimulus persons exceeding those ages are judged differently depending upon whether they are male or female. Photos of older females are judged consistently younger by the elderly subjects, whereas the age of the judge does not seem to have a systematic effect on the ages attributed to the photos of older males.

Determinants of Risk Taking for Others
(NICHHD, 784-4)

N. Kogan
A. Teger

PURPOSE: The present research represents an extension of earlier work by Zaleska and Kogan (RB-69-9) on the comparison of individual and group decision-making for other people. The study examines the effects of friendship and reciprocity of decision-making between self and other on preferred risk levels.

PROGRESS: A report (RB-71-56) has been issued and submitted for publication. A condensed version was published in the *Proceedings of the 79th Annual Convention, APA*, 1971, 247-248, and a more extended version is in press in the *British Journal of Social and Clinical Psychology*.

MAJOR FINDINGS: Decision-making for others was not affected by extent of friendship between the decision-maker and the other. On the other hand, a significant reciprocity effect was obtained. In deciding for another, subjects were inclined to take greater risk under nonreciprocal conditions than in the case where the other would in turn be deciding for the decision-maker herself.

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts (NICHHD, 784-4) N. Kogan

PURPOSE: In the last 15 years there has been a growing body of research into motivational factors in achievement in both the classroom and other situations. Such research has been possible because of the development of useful measures of achievement motivation. Although psychologists have long recognized the influence of the interpersonal context upon the individual's striving and performance, there has been little research on the specific effects of such contexts. The principal aim of this study is to examine the impact of two commonly occurring contexts--adult-evaluative and peer-competitive--on the achievement motivation and performance of elementary school boys.

PROGRESS: A report (RB-71-49) has been issued, and a version for publication is currently in preparation.

MAJOR FINDINGS: On the whole, the results in both the peer-competitive and adult-evaluative contexts conformed to Atkinson's motivational model of risk taking--i.e., a preference for intermediate risk taking in "hope for success" subjects and a preference for more extreme risk or caution in "fear of failure" subjects. Of further interest was evidence showing significant variation in risk levels across contexts for high defensive boys--more cautious under "adult evaluation" and more risky under "peer competition"--and the absence of such variation for low defensive boys. (Project conducted in collaboration with Mr. J. Damm of the University of Queensland, Australia.)

Cognitive Styles and Attitudinal Consistency

S. Messick

(NICHHD, 784-9)

PURPOSE: Some of the major concerns of social psychology in recent years have dealt with the dynamics of attitudinal consistency--in particular, (a) with the extent to which both cognitive and affective components of attitude systems tend to be organized in psychologically consistent patterns, (b) with the operation of motivational pressures toward consistency, and (c) with the specific mechanisms whereby inconsistency is resolved. The present study investigates the importance of reliable individual differences in the tendency toward attitudinal consistency, the generality of their operation across different measurement approaches and different response processes (such as interpersonal perception and opinion change), and the extent of their relation to other cognitive and personality variables.

PROGRESS: Attitudinal consistency was measured using four experimental procedures that had previously formed the basis for four major studies in this area: The McGuire Logical Consistency Test, the Morrissette Balance Completion Test, the Osgood and Tannenbaum Congruity Test, and the Rosenberg and Abelson Fenwick Test.

These four experimental procedures were administered, along with a measure of intolerance of attitudinal inconsistency, to a sample of 200 high school seniors who had been extensively tested earlier with a battery of cognitive and personality measures. One ancillary function of the present study will be an evaluation of the replicability with high school students of the mean effects observed in the four previous experiments by Morrissette, McGuire, Rosenberg and Abelson, and Tannenbaum. The major emphasis, however, is upon the consistency of interrelations among these four procedures, upon the number and nature of dimensions required to account for individual differences in attitudinal consistency, and upon the cognitive and personality correlates of those dimensions. Particular attention will be paid to the relations between categorizing styles and attitudinal consistency, since level of conceptual differentiation and cognitive complexity may moderate consistency pressures. (Study done in collaboration with J. E. Singer of State University of New York at Stony Brook.)

Social Stratification (NICHHD, 784-16) L. Stricker

PURPOSE: A vast amount of research has been devoted to the identification and development of objective indexes of social stratification, such as education and occupation. Although this research has been based almost entirely on samples of whites, these indexes have often been used to measure the stratification of Negroes, despite suggestions that these indexes are not comparable for the two groups.

An inherent problem with these indexes, even when they are restricted to whites, concerns the identification of the cutting points for differentiating between the social classes. Empirical attempts to identify these points have largely been unsuccessful and a theoretical controversy continues over whether the social classes are distinct entities, reflecting discontinuities in stratification, or are arbitrary classifications that have been imposed on a continuum.

This study has two aims: (a) to identify the dimensions of social stratification for both Negroes and whites, and the indexes that best reflect these dimensions, and (b) to determine whether distinct social classes exist and, if so, whether they can be empirically differentiated.

PROGRESS: A comprehensive, highly structured interview schedule was prepared, focused on information that directly reflects the major aspects of social stratification or are important correlates of it. The schedule was used in household interviews with representative samples of Negro and white household heads in a typical northern city. The data are being analyzed.

Effects of Models on Creative Performance in Relation to Anxiety, Ability, and Sex (NICHHD, 786-10)

N. Frederiksen
F. Evans

PURPOSE: This study involves an attempt to improve performance on the Formulating Hypotheses test (which requires subjects to develop hypotheses to account for data) by presenting models of "acceptable" performance that stress either quantity or quality of responses. Results of a previous study (Journal of Educational Psychology, 1969, 60, 465-475) that involved only the "quantity" models showed that (1) the use of models increased the number of hypotheses formulated; (2) the training did not transfer to a test of divergent production (Guilford's Consequences test); (3) test anxiety was curvilinearly related to performance; and (4) there was a significant interaction of anxiety and verbal ability. The purposes of the present study were to attempt to replicate the previous findings and also to investigate (1) the effects of models emphasizing quality as well as the quantity models, (2) sex differences, and (3) the relationships of ideational fluency and verbal ability to performance.

PROGRESS: Data were obtained from about 400 college freshmen in two eastern state colleges. New methods of scoring were developed to yield scores reflecting quality as well as quantity. A multivariate analysis of variance was completed. A report of the study has been published (Effects of Models of Creative Performance on Ability to Formulate Hypotheses. Journal of Educational Psychology, 1974, 66, 67-82).

MAJOR FINDINGS: Both the quantity and quality models were effective in modifying behavior in the expected direction. Ideational fluency was related to the number of hypotheses, and verbal ability was related to scores reflecting quality. Females were on the average superior to males with respect to scores reflecting quantity of responses. Test anxiety was not significantly associated with performance. Weak evidence of treatment-anxiety and sex-vocabulary interactions were found.

Developmental Changes in Self/World Differentiation (NICHHD, 787-30)

A. H. Baker

According to certain leading developmental theorists (e.g., Piaget, Werner, Schachtel), boundaries between self and world are lacking in human infancy and evolve only with higher levels of development. In particular, it is assumed that the relationship between self and world can vary from one of nondifferentiation (i.e., fusion or linkage) which prevails in infancy to one of relative differentiation (i.e., separateness, distance from, or oppositeness) which characterizes the adult. Several interrelated studies have attempted to demonstrate that even among normal adults, degree of self/world differentiation can vary, and that such variation has measurable effects upon perception. We have recently extended this work (1) by studying age changes in perceptual phenomena, using children as subjects; and (2) by studying individual differences in perceptual experience among adult subjects along one or more dimensions of self/world differentiation.

Study 1

PURPOSE: In studying age changes in perceptual phenomena, the specific dependent variable chosen for study is the psychophysical error of anticipation (= undershooting) which has long been known to occur in space localization experiments when some variable stimulus, presented in an ascending or descending trial, is moved toward some standard value. The central hypothesis which shaped our earlier studies with adults--viz., that with increase in degree of self/world differentiation there is an increase in magnitude of undershooting--was derived in part from developmental theory. As such, it would seem a most natural extension of this work to explore directly whether there are changes during ontogenesis in the effects of these variables which the earlier research demonstrated affects adult performance. It is thus proposed to undertake a series of studies, cross-sectional in design, using children between the ages of 7 and 18, while at the same time extending our work with adult subjects.

PROGRESS: Progress has been achieved to date in the following areas: (1) a laboratory for conducting the planned perceptual-cognitive experiments has been built; (2) necessary equipment has been constructed; (3) pertinent literature review has been accomplished; (4) a well qualified research assistant, with some expertise in the area of perception, was hired and trained in a variety of dark-room laboratory procedures which are necessary for conducting the planned research; and (5) data from some published studies by other authors have been procured and analyses have been undertaken to clarify issues relevant to this project.

The most widely supported finding emerging from this series of interrelated studies is that there is a systematic shift from undershooting error (stopping too soon) toward overshooting error (letting the stimulus go too far) over the course of repeated trials. This shift phenomenon occurs both when the experimenter and when the subject controls the stimulus; both when the stimulus is moved continuous or in steps; both with the psychophysical methods of limits and of adjustment. It occurred over the course of several short sessions (where level of fatigue is presumably constant). We have interpreted this phenomenon as reflecting a shift from an active (= greater arousal) to a passive (= lesser arousal) orientation due to task repetitiveness. For the types of situations employed, in which the task was externally imposed and the stimulation repetitive, it was consequently hypothesized--in the context of a test of this interpretation--that the younger the child, the quicker the tendency to become bored and thus the more rapid occurrence of the shift phenomenon. This hypothesis was recently confirmed, using children ranging from 6 to 18 years of age.

Preliminary pilot studies have raised the possibility that (1) degree of undershooting/overshooting error (independent of the shift phenomenon) and/or (2) the shift phenomenon itself may constitute individual-differences variables reflecting underlying perceptual-style and personality characteristics. To explore this possibility, an extensive battery of tests has been devised which provides multiple measures of both undershooting errors and of shift in such errors over repeated trials. This battery, coupled with many potentially fruitful personality and cognitive style measures, has been administered to 120 individuals. Data analyses are now in progress. In this context, a procedure for obtaining both types of errors in a group administered setting has been devised.

Study 2

PURPOSE: The present study has as its focus individual differences in the degree of self/world differentiation among normal adult subjects. Although a number of studies in the past decade have demonstrated that variation in the relationship between self and stimulus object does have measurable effects upon perceptual experience, for the most part such studies have been experimental in nature--that is, conditions were varied, and the goal was to demonstrate that some general effect resulted from this variation. Studies of individual differences in the ongoing relationship between self and object have been quite rare and almost always limited to a single perceptual situation. The present research seeks to go beyond this by attempting--for the first time--to explore whether there are consistent individual differences in self/world orientation which can be assessed by a battery of tests tapping a variety of perceptual and cognitive situations.

PROGRESS: Progress has been achieved to date in the following areas: (1) space has been procured and a laboratory built at the Psychology Department of Florida State University for conducting the planned perceptual-cognitive experiments; (2) all equipment has been designed and is now in operational use; (3) plans have been finalized regarding the initial phases of this project; (4) an extensive set of materials was prepared specifying in detail the procedures, instructions, ways of recording data, etc. for each of the tasks decided upon; (5) the person who will be collecting the data during the initial phases was trained in the appropriate laboratory procedures; (6) to insure that standard dark-room lab procedures would obtain, the present researcher spent two days at the Florida State University laboratory checking out all of the procedures, and (7) an initial study involving part of the total proposed battery has been conducted. This involved the standard rod-and-frame test; a similar test in which a rod alone is presented to the subject; two variants of the rod-and-frame test--one with instructions which make a person highly aware of his own body during the task, and another in which the person is asked to suppress all information from his body while carrying out the task (these latter two based on some research by Baker, 1968); and a novel task, using the rod-and-frame apparatus, in which the subject is asked to line up the stimulus rod with the longitudinal axis of his tilted body (after Baker, 1963).

Results of some of the above findings have been or will be reported as follows:

(1) Habituation (overshooting) and anticipation (undershooting) errors in visual and tactual space localization. Presented at meetings of The Psychonomic Society, November 1969, St. Louis, Missouri.
(2) Effects of stimulus structure, mode of inspection, and repeated trials upon anticipation-habituation errors in space localization. Presented at the meetings of the American Psychological Association, September 1970, Miami, Florida (see also ETS RB 70-45).
(3) Changes in body perception during repeated trials. Presented at the meetings of the American Psychological Association, September 1970, Miami, Florida (with Jerry Schlater and Seymour Wapner).
(4) Shift from undershooting toward overshooting error over repeated trials with a body adjustment task. Presented at meetings of the Psychonomic Society, November 1970, San Antonio, Texas (with Theodore Leventhal).
(5) Psychophysical error of anticipation and the method of limits: A developmental analysis. Presented at meetings of the International Society for the Study of Behavioral Development, July 1971, Nijmegen, The Netherlands.
(6) Shift from undershooting toward overshooting error over repeated trials: A developmental analysis. To be presented at meetings of the International Society for the Study of Behavioral Development, August 1973, Ann Arbor, Michigan (with Laraine M. Schwartz and Brian Mishara).
(7) Toward formalizing the "undershooting to overshooting shift" as a general phenomenon: Findings with subject-stimulus location fixed. To be presented at the Fourteenth Annual Meeting of the Psychonomic Society, November 1973, St. Louis, Missouri (with Laraine M. Schwartz and Brian Mishara).

(Study done in collaboration with Seymour Wapner from Clark University and Joseph Grosslight of Florida State University.)

International Study of Educational Attainment in French as a Foreign Language (USOE through Teachers College, Columbia University, 798-19)

J. Carroll

PURPOSE: This is a part of the International Study of Educational Attainment (IEA) being conducted by an International Council headed by Professor Torsten Husén of the University of Stockholm, and in which achievement tests in English as a Foreign Language, French as a Foreign Language, Civics, Literature, Reading Comprehension, and Science were administered in a number of countries for the purpose of comparative studies. The French as a Foreign Language tests were monitored by an international committee of which Mr. Carroll is chairman and Mr. Clark is a member. Tests of listening, speaking, reading, and writing were prepared for various levels of achievement, ranging from that of a 10-year-old population to that of the pre-university level.

PROGRESS: The tests approved by the International Council of the IEA Study in December 1969, with minor modifications, were administered in a number of countries in the spring of 1971 to large, statistically controlled samples of students in grades 10 and 12. A set of tests was also administered to 10-year-old pupils in the United States. The administration of the tests was handled by national centers in the various countries. Major statistical analyses were accomplished partly at Teachers College, Columbia University, under the direction of Richard Wolf and partly at the University of Stockholm under the direction of T. N. Postlethwaite and others.

During 1973-74 at ETS, Mr. Carroll conducted further statistical analyses of the data and prepared a preliminary report of the major findings for a conference on all the IEA results that took place at Harvard University, November 11-16, 1973. In comparing various countries and populations with respect to overall performance, it was necessary to establish common scales by vertically equating the Population 2 and 4 tests. Sources of variation in the proficiency test scores were studied in six classifications: (1) Background factors (socioeconomic status, sex); (2) Type of student program; (3) Time factors (age, grade, number of years of French study, age at start of French study); (4) Conditions of instruction (teacher competence, classroom teaching methods); (5) Kindred student variables (student interest, aptitude, outside activities, amount of study time, etc.); and (6) student verbal ability as measured by a word knowledge test given in the various mother tongues. Generally, there were significant amounts of predictor variance contributed by each of these classifications. In different countries and populations, multiple correlations of criterion variables with predictor variables ranged from .52 to .89. Of particular interest was the fact that considerable amounts of variance could be traced to school factors such as the amount of time devoted to French study, teacher competence in French, and classroom teaching methods.

Because of delay in the completion of further statistical analyses, it was not possible to prepare the final report during 1973-74, but Mr. Carroll plans to complete this report during the summer of 1974. The published report will parallel the books that became available in May 1973 concerning other subject matters (International Studies in Evaluation. Stockholm:

Almqvist & Wiksell, and New York: Halsted Press Division of John Wiley. I: Science Education in Nineteen Countries, by L. C. Comber and J. P. Keeves; II: Literature Education in Ten Countries, by A. C. Purves; III: Reading Comprehension Education in Fifteen Countries, by R. L. Thornkike.)

A report on one major aspect of the study--the development of the French proficiency tests, along with an analysis of their measurement characteristics--was prepared in the form of a Research Bulletin (RB-74-15), "Language proficiency tests developed for the IEA International Study of Achievement in French as a Foreign Language." This report will eventually appear as a chapter in a book, Current Trends in Language Testing, edited by B. Spolsky, to be published by Mouton, The Hague.

Peer Group Interaction in Infancy

(NIE, 798-62)

M. Lewis
J. Brooks
L. Michalson
G. Young

PURPOSE: A series of studies has been designed to explore infants' relationships to other infants. There is reason to believe that infants react differently in the presence of children and adults, just as they respond differently to mothers and strangers. Infants may also form social relationships with persons other than familiar adults. Therefore, we are examining the development of and interaction within peer groups.

PROGRESS AND MAJOR FINDINGS: (a) Social behavior in infant play groups. Eight groups of four 12-month-old infants and their mothers, and eight groups of four 18-month-old infants and their mothers were observed during free play. Each group consisted of two boys and two girls. The infants were allowed to respond naturally to their mothers, to unfamiliar peers, and the peers' mothers.

The results show that children as young as one year of age do respond differently to their mothers and the strangers. They stay close to and touch their mothers while looking frequently at the strangers. More importantly, the infants respond differently to the unfamiliar adults and children. They look at, smile at, and play more with the other children than with either their mothers or the other mothers. Moreover, the infants do not interact randomly with their peers, but rather the girls interact more with girls and the boys with boys.

The findings indicate that social behavior originates at very young ages. Infants as young as one year of age seek out and respond in different ways to people around them. A complete report of the findings will be presented at a Conference on The origins of behavior: Friendship and peer relations, and will appear in a book by the same title edited by M. Lewis and published by Wiley.

(b) Friendship and peer relationships. Infants have also been seen in same-sex dyads with familiar and unfamiliar peers. We are interested in whether infants will respond differently and will prefer interacting with the infants who are "friends" rather than strangers.

(c) The development of peer relationships. The development of infant relationships is being studied. Infant dyads and their mothers are observed together over a three-week period to see if relations and behavior change over time and to explore the origins of friendship.

Social Perception in Infancy
(NIE, 798-62; NIMH, 828-1)

J. Brooks
M. Lewis

PURPOSE: Infants' responses to persons have traditionally been studied under the rubric of attachment and fear, and has involved the ability to differentiate between mother and other. However, infants also discriminate between other familiar persons (i.e., father, siblings) as well as between classes of strange persons. In the present series of studies, we are exploring the various characteristics of persons which are used by the infant in categorizing his social world. These physical characteristics, gender, height, and facial configuration as well as the familiarity-unfamiliarity dimension, have been systematically varied.

PROGRESS AND MAJOR FINDINGS: (a) Self, other and fear: the reaction of infants. Infants between eight and 18 months of age were introduced to five different social events: strange adult male and female, strange four-year-old female, mother, and self. The infants' responses indicated that approach affects stimulus differentiation (in terms of fear), while age affects level of response intensity. Strangeness of the social event was not sufficient to explain the results, and ethological and cognitive theories were discussed. It was suggested that the cognitive construct of self be introduced to explain the results. It appears reasonable to assume that by eight months the self as differentiated from other is well established. Moreover, we propose that some specific categories of self, such as size, sex, or efficacy, may be established as well. The paper was presented at the Eastern Psychological Association meetings and is available as RB-72-23. A more complete version was presented at a Conference on The origins of behavior: Fear, Princeton, N. J., October 1973, and will appear in M. Lewis & L. Rosenblum (Eds.), Fear: The origins of behavior, volume II. New York: Wiley, in press.

(b) Infants' social responses to social events. To test whether size or facial configuration accounted for the positive responses to children and negative ones to adults in the preceding study, these two characteristics were systematically varied. Three different classes of strangers--child, child-sized adult, and normal-sized adult--approached 40 infants. The infants' responses indicated that the small adult was not seen as a child or a normal-sized adult. This suggests that both features are utilized in an interdependent fashion. We hypothesize that facial configuration controls the direction of affect, while size controls the intensity.

Study of Bachelor's Degrees J. Warren
(Western College Association, 801-9)

PURPOSE: As college degrees become more important as a social and occupational credential, their meaning and substance become correspondingly more critical. A series of meetings with about a dozen representatives of member colleges of the Western College Association has been held to examine issues in the meaning of degrees that would be useful topics for a detailed study.

PROGRESS: The committee sketched out a domain of potential studies, and 55 representatives of WCA institutions rated the importance of 15 issues the committee had identified. A study to identify the range and variety of competencies associated with degrees in nine fields of study at 90 colleges and universities in California will be initiated in September 1974.

The Effect of Time on Attachment as Measured in a Free Play Situation (Spencer, 815-1)

J. Brooks
M. Lewis

PURPOSE: Attachment studies vary in terms of various situational dimensions. One of these dimensions is amount of time the mother-infant dyad is observed. Length of time spent in the playroom may facilitate or inhibit attachment behavior as exhibited by the infant. In fact, length of time may be a more important dimension than number of toys or strangers present. Our pupose was to investigate the effects of time on the infant's behavior.

PROGRESS AND MAJOR FINDINGS: Eighty-four 12-month-old infants were observed for 15 minutes in a low-stress free play situation. The amount of looking at, touching, vocalizing to, and remaining in proximity to the mother were analyzed for five 3-minute time periods. Three of these behaviors--look, touch, and proximity--increased over time. Thus, time spent together influences attachment between mother and child independent of any manipulation has been demonstrated. This suggests that other ecological variables must also be controlled or explored.

A paper appeared in Child Development, 1974, 45, 311-316.

Mother-Infant Interaction
(Spencer Foundation, 815-1)

M. Lewis
S. Lee-Painter

PURPOSE: To explore the ongoing interaction between the mother and her 12-week-old infant, and to follow through and investigate this communication as the infant develops through its first and second years of life.

PROGRESS: Ninety-seven 12-week-olds of both sexes and from 5 social classes have been seen with their mothers in their own homes. The data collected include a wide variety of maternal and infant behaviors. Emphasis has been placed on observing the chains of interaction that form the communication network of the mother and her three-month-old infant. These 12-week-old infants have been brought to the laboratory by their mothers after the home visit. They have been given measures of visual and auditory attention, which have been shown to be associated with (related to) cognitive development. At the same time, they have also been given some of the standard measures of development, such as the Bayley scales. The data from this phase of the research are being analyzed and some preliminary results are discussed below.

When the same children reached the age of one year, they were again brought to the laboratory and were observed in a standardized free-play situation. Observation methods included a dictated continuous behavioral account of the infants' activities, and the recording of four measures of infant attachment and behavior. Data analysis of this phase of the research is in progress, and we soon will be able to make some definite statements about the relationship between mother-infant interaction as seen at 12 weeks and the attachment bond of the one-year-old.

In the third phase of the project, the infants returned to the laboratory around their second birthday for a final, two-session series of observations. The standardized free-play situation (modified to meet the needs of a two-year-old) is again employed, as well as several cognitive measures which we hope will lead us to some conclusions about the relationship of the nature of mother-infant interaction and cognitive development. The videotapes which were made of segments of the two-year visit have been reviewed in terms of separation behavior and the data from these tapes are discussed elsewhere (Lewis & Weinraub).

MAJOR FINDINGS: Some preliminary data from the 3-month phase of the project have been analyzed. Some of the major results have been the contribution of models by which data from this and other research can be made meaningful. Specifically, the model of a general communication matrix is one which takes into account the diverse patterns of mother-infant interaction which are the essence of their communication network. From this model we come to see that it is not only the frequency of behavior with which we must be concerned, but the chaining effect of initiation and response of these behaviors as well. Thus, for example, mothers of boys smile more than mothers of girls in terms of behavior frequency. However, when maternal smiling is looked at in terms of interaction, i.e., when it is an initiation of or a response to an infant behavior, mothers of girls smile more than mothers of boys. The development of a Markov model for analyzing the vocalization chains of the mother and infant rises naturally from the general communication matrix model. By use of the Markov model, we can begin to detect the fine lines of communication between the mother and her infant. Several papers have reported on these and other preliminary findings: Lewis, M. State as an infant-environment interaction: An analysis of mother-infant interaction as a function of sex. Merrill-Palmer Quarterly, 1972, 18, 95-121; Lewis, M., and Freedle, R. Mother-infant dyad: The cradle of meaning. Research Bulletin 72-22. Princeton, N. J.: Educational Testing Service, 1972; Lewis, M., and Lee-Painter, S. An infant's interaction with its social world: The origin of meaning. Paper presented at the Canadian Psychological Association meetings, Montreal, June 1972; Lewis, M., & Lee-Painter, S. An interactional approach to the mother-infant dyad. To appear in M. Lewis and L. Rosenblum (Eds.), The origins of human behavior. New York: John Wiley & Sons, 1973, in press.

In addition to these findings, we have begun to intercorrelate various parts of the 12-week observation, specifically, maternal behavior frequency and three measures of the infant's cognitive development: (1) the Mental Development Index of the Bayley Scales; (2) a Piagetian instrument of sensori-motor development--the Escalona-Corman Object Permanence Scales; and (3) the measure of attentional distribution-response decrement to a redundant signal. Two major findings resulted from this analysis: (a) no relationship was found between the various measures of cognitive development; (b) a negative relationship was found between maternal behavior frequency and the cognitive scores. A paper recently presented at the Eastern Psychological Association meetings (Mother-Infant Interaction and Cognitive Development, Susan Lee-Painter & M. Lewis) discusses the results in terms of the validity of infant development measures and the impact of the mother-infant relationship.

The Scales used to collect the mother-infant interaction data have been revised and are now in printed form along with an explanation of the data collection process and the various data analyses available. (Mother-Infant Interaction Scales, Michael Lewis and Susan Lee-Painter, ETS, 1974.)

Checking Procedures for Estimating Item and Person (Ability) Parameters by Analyzing Artificial Data

F. Lord

(NSF, 816-1)

PURPOSE: To check the accuracy of estimates of item-characteristic curve parameters and examinee ability parameters obtained by current methods.

PROGRESS: One set of artificial data resembling the SAT-Verbal test was generated and analyzed.

MAJOR FINDINGS: Good agreement was found between estimates and true values, except for the pseudochance-level parameters for easy items which do not affect the fit of the model to the data.

Design and Evaluation of Novel Tests for Practical Use (NSF, 816-1; CEEB, 201-20)

F. Lord
M. Wingersky

PURPOSE: To use available estimated ability and item-characteristic-curve parameters to predict the effect of a variety of proposed changes in the SAT.

PROGRESS: A computer program has been written by Marilyn Wingersky that gives the relative efficiency of variously modified versions of the SAT compared to the present test. Various designs for a multilevel SAT have been tried out on this program, as well as designs for a shortened version of the SAT.

Formula Scoring and Number-Right Scoring (NSF, 816-1) F. Lord

The assumption that examinees either know the answer to a test item or else guess at random is usually totally implausible. A plausible assumption is outlined, under which formula scoring is found to be clearly superior to number-right scoring. A suitable research strategy is discussed. (RB-74-1)

Individualized Testing and Item Characteristic Curve Theory (NSF, 816-1) F. Lord

An elementary survey of item characteristic curve theory is presented, centered around the problems of individualized ("tailored") testing.

To appear in R. C. Atkinson, D. H. Krantz, R. D. Luce, and P. Suppes (Eds.), Contemporary developments in mathematical psychology. San Francisco, Calif.: Freeman. (RB-72-50)

Quick Estimates of the Relative Efficiency of Two Tests as a Function of Ability Level (NSF, 816-1)

F. Lord

When comparing two tests that measure the same trait, an overall comparison is not enough. Separate comparisons should be made at different levels of the trait. A simple, practical, approximate formula is given for doing this. The adequacy of the approximation is illustrated using data comparing seven nationally known sixth-grade reading tests. (RB-73-67)

The Relative Efficiency of Two Tests as a Function of Ability Level (NSF, 816-1) F. Lord

A new formula is developed for the relative efficiency of two tests measuring the same trait. The formula expresses relative efficiency solely in terms of the standard errors of measurement, and, surprisingly, the frequency distributions of true scores. Approximate methods for estimating relative efficiency may make this function routinely available. A numerical illustration compares new and old estimates of relative efficiency for subtests from the Scholastic Aptitude Test. (RB-73-41)

Variance Stabilizing Transformation of the Stepped-Up Reliability Coefficient (NSF, 816-1) F. Lord

The stepped-up reliability coefficient does not have the same standard error as an ordinary correlation coefficient. Fisher's z -transformation should not be applied to it. Appropriate procedures are suggested.

Published in the Journal of Educational Measurement, Spring 1974. (RB-72-48)

The 'Ability' Scale in Item Characteristic Curve Theory F. Lord
(NSF, 816-2)

For the six available sets of empirical data, the discrimination (slope) parameter of the logistic item characteristic curve was found to have a significant positive correlation over items with the difficulty (location) parameter. This unpleasant situation can be eliminated by a suitably chosen transformation of the ability scale. An example of the transformation is given and some empirical results of using the transformed scale are shown. (RB-74-19)

Relative Efficiency of Number-Right and Formula Scores F. Lord
(NSF, 816-2)

A formula is derived for estimating at any specified ability level the relative efficiency of number-right score and formula score. Analysis of one set of data is summarized. (RB-74-9)

Assessment of Infant Socioemotional Development (OCD, 817-2)

M. Lewis
L. Michalson

PURPOSE: To finish constructing and to refine the Infant Scales of Socioemotional Development and to use this instrument to assess socioemotional development in several kinds of infant intervention programs. Although the primary use of the instrument is to evaluate the development of infants between 3 and 24 months old in a day care environment, eventually it will be modified for the purpose of exploring the effects of other kinds of infant intervention programs.

PROGRESS: The Scales consist of a behavior-by-situation matrix which is divided into five socioemotional domains: fear, anger/frustration, happiness, attachment, and competence. Sixteen persons in the Research Division were asked to rate each of the behaviors on a 1-5 scale according to the intensity of emotion expressed by each behavior. On the basis of the means of the judges' ratings the behaviors within each domain were assigned an intensity rating from 1-5.

To date 32 infants have been observed in a day care center, and their behaviors in a variety of situations have been recorded. Based on these data reliabilities of the domain scores were computed; alpha coefficients were above .90 for the items in each domain. In addition, principal components analyses were carried out on the data of each domain, and the results suggest that each socioemotional domain may be composed of two or three subcomponents.

Further investigations will be directed toward gathering normative data, examining the relationships among the domains, and establishing the developmental course of these socioemotional variables through the first years of life. In addition to observing day care infants, a comparable group of home-care infants will be rated on the assessment instrument in an effort to evaluate the effects of the intervention program.

Developmental Studies of Selective Attention (NICHHD, 826-2)

G. Hale

PURPOSE: This research addresses the question, to what extent do children attend selectively to a single aspect of stimulus materials and ignore other, extraneous information? To identify developmental changes in this process from early school age to preadolescence, two general approaches to the study of attention are combined: (a) assessment of learning and recall of central and incidental aspects of stimuli and (b) examination of the distracting effects of extraneous stimulation.

PROGRESS: The following publications issued during this fiscal year describe earlier studies in this research program: (1) Hale, G. A., & Taweel, S. S. Children's component selection with varying degrees of training. *Journal of Experimental Child Psychology*, 1974, *17*, 229-241; (2) Hale, G. A., & Lipps, L. E. Stimulus matching and component selection: Alternative approaches to measuring children's attention to stimulus components. *Child Development*, 1974, *45*, 383-388; (3) Hagen, J. W., & Hale, G. A. The development of attention in children. In A. D. Pick (Ed.), *Minnesota Symposia on Child Psychology*, Vol. 7. Minneapolis: University of Minnesota Press, 1973. Pp. 117-140.

The following papers, in press during this fiscal year, were also based on this research: (1) Hale, G. A., & Piper, R. A. The effect of pictorial integration on children's incidental learning. *Developmental Psychology*, in press; (2) Hale, G. A., & Stevenson, E. E. The effects of auditory and visual distractors on children's performance in a short-term memory task. *Journal of Experimental Child Psychology*, in press; (3) Hale, G. A., & Taweel, S. S. Age differences in children's performance on measures of component selection and incidental learning. *Journal of Experimental Child Psychology*, in press. Two additional papers are derived from this project: (1) Hale, G. A., Taweel, S. S., Green, R. Z., & Flaugher, J. Flexibility of children's attention: Instructions regarding deployment of attention in a component selection task. Educational Testing Service Research Bulletin, in press; (2) Hale, G. A., & Green, R. Z. Children's component selection as a function of stimulus integration and relative salience of components. Educational Testing Service Research Bulletin, in review.

Data have been collected for six experiments during the current fiscal year--three studies of incidental learning and three studies of distractibility. Preliminary analyses have been conducted for all but one experiment, and the findings are summarized below.

MAJOR FINDINGS: The first experiment involved a total of 280 children at ages 9 and 12, and the second study, a replication and extension of the first, included 240 children at ages 9 and 12. The children were given a series of learning trials with compound pictorial stimuli and were required to attend to a single element in each stimulus. An incidental recall test then determined the child's memory for information about other elements. The children were given either 6 or 12 seconds of exposure to the materials on each learning trial, and either 8 or 16 learning trials. Doubling the exposure time per trial increased incidental recall at age 9 but not at age 12; however, it improved performance on the learning trials at age 12 more than at age 9. As children approach adolescence they apparently make more efficient use of additional exposure time; they use the time to improve learning, while younger children attend to nonessential aspects of the material. Increasing the number of learning trials, however, did not affect 9- and 12-year-olds differentially. Thus, these two variables appear to represent functionally different means of increasing stimulus exposure. In the third experiment these same variables were examined with 280 children at ages 5 and 9, using a simplified task. Increasing stimulus exposure (via both methods) facilitated recall of incidental information to a greater degree at age 9 than at age 5. An increase with age in children's ability to take in and recall visual information appears to be a major determinant of incidental learning prior to middle childhood. However, a developmental increase in children's use of selective attention contributes to declining incidental learning from middle childhood to adolescence. Other results were: (a) the amount of incidental recall depended on whether the test followed immediately after learning or after a delay, (b) recognition and recall tests of incidental memory produced comparable patterns of results, and (c) the presence of incidental stimulus features distracted the children as they performed the learning task.

In the first distraction study, 192 children at ages 5 and 9 years performed a learning task in the presence or absence of auditory and visual distractors. Visual stimulation was more distracting for 5-year-olds than 9-year-olds, while auditory stimulation proved more distracting for 9-year-olds. It is said that as children mature they shift from reliance on visual imagery to use of a verbal memory. The present results suggest that the more potent distractor at a given age level is one that interferes with the dominant mode of information storage. Nonmeaningful visual stimulation (patterns flashing in periphery of visual field) had a more distracting effect at age 5 than at age 9, while meaningful visual distractors (flashing pictures) had an equally distracting effect across age levels. Apparently, distractors designed to capture the child's interest are functionally different from distractors that simply "intrude" upon performance. Other results were: (a) nonmeaningful visual distractors superimposed on the learning materials maintained their distracting effect across ages, perhaps because they could not be ignored by focusing on a particular spatial location, and (b) steady noise, presumably not sufficiently "intrusive" to interfere with use of verbal memory, had a declining effect across ages.

The second experiment looked at the effects of these distractors with 192 children at ages 9 and 12, using a more difficult task. The most important result was that the effects of distraction were greater at age 12 than at age 9. Because they are adapting to a new school environment (junior high school) and entering adolescence, children around age 12 may be particularly susceptible to distraction. In the third experiment, with 144 children at ages 5 and 9 years, distraction effects were assessed with tasks varying in difficulty; data have been collected but not analyzed as yet. In general, these experiments underscore the inadequacy of the hypothesis that children's response to distraction simply decreases with age. Rather, the developmental effects of a distractor depend on its interest value, its relation to the processes involved in task performance, and other factors. Furthermore, distractibility may not decline linearly with age but may undergo upswings during certain age periods.

Development of Interest Measurement Instruments for Assignment of Army Officers to Appropriate Positions

(Army Research Institute, 827-1)

J. Campbell
R. Boldt
R. Flaugher
D. Rock

PURPOSE: This project is designed to provide the Department of the Army with dimensions of Army Officer interest appropriate for use with a computer-aided career counseling system, a set of instruments to measure these dimensions, and with similarity measurement and clustering techniques for grouping officers and positions.

PROGRESS: In the first phase of the project, forms were developed with approximately 200 items descriptive of Army job characteristics. Approximately 250 officers were asked to indicate how well each of these items described their present position, the assignment they would prefer next, and an assignment they would not like to have. The immediate superiors of these officers were also asked to indicate how well the items described the officers' present position. Factor analyses of these data yielded seven factors which were labeled (1) troop duty in the field, (2) troop duty in garrison, (3) hazardous duty, (4) supply activities, (5) staff activities, (6) relations with civilians, and (7) instructional activities. In the second phase, new forms will be developed to try to elicit a different set of dimensions.

Infant's Responses to Pictures of Mother, Self, and Other (NIMH, 828-1)

J. Brooks
M. Lewis

PURPOSE: This study was designed to explore further the infants' ability to discriminate between familiar and unfamiliar persons and to differentiate among strangers. We recently found that infants exhibit negative affect (fear and wariness) to adult strangers and positive affect to a child stranger (Lewis & Brooks, 1974). In addition, infants responded quite positively to their mirror image as well as to their mothers. The present study extends our inquiry into the infant's ability to recognize himself as well as differentiate among various strangers when presented two-dimensional representations.

PROGRESS AND MAJOR FINDINGS: Pictures of the infant's mother, the infant himself, other infants, 5- and 10-year-old children, and adults were shown to 54 infants in four age groups: 10, 12, 16, and 18 months. Pictures of both male and female strangers were presented for the baby, children, and adult stranger conditions. In general, the older infants were more likely to differentiate among the pictures than were the younger infants.

The older infants discriminated between pictures of familiar and unfamiliar persons as well as between classes of strangers. Specifically, they responded differentially to mother and adult stranger, self and baby stranger, and baby and adult stranger conditions. In addition, infants preferred looking at same sex rather than opposite-sex peers. The concept of self, gender identity, and social perception are discussed in terms of the data.

A paper was presented at the Eastern Psychological Association meetings in Philadelphia, March 1974. A more complete version will appear in P. Salapatek & L. Cohen (Eds.), Perception in infancy, Academic Press, in press.

Visual Self Recognition in Infancy (NIMH, 828-1)

J. Brooks
M. Lewis

PURPOSE: Since there is little research on the concept of self in children under two years of age, we are conducting exploratory research on the developmental trends in self recognition. Infants' reactions to pictures of themselves, responses to a mirror, and reactions to viewing themselves on a TV monitor are being measured in order to discover the dimensions underlying self recognition. In addition, the relationship of this competency to other abilities, both social and cognitive, are being studied. The relationship of early self recognition and self concept in the preschooler is also being explored. If antecedents are defined, intervention may be possible so that a child's self concept may be altered in a positive way.

PROGRESS AND MAJOR FINDINGS: A series of studies on this topic have been designed, and data collection has been completed for two.

System of Interactive Guidance and Information (Carnegie, 842-3)

M. Katz

PURPOSE: A computer-assisted system of interactive guidance and information is being developed to improve career decision-making by students in junior colleges.

SIGI is a guidance system which allows the student to interact with a computer in such a way as to examine his own values, obtain and use relevant information, interpret predictive data, and formulate plans. This interaction assists the student in arriving at tentative career decisions and in modifying them as he gains new insights and additional information. The decisions involve both educational and occupational options. Emphasis, however, is not merely on the content of decisions but on the process of decision-making. As the student progresses through SIGI, he learns to move freely within the structure of the system. In gaining control of the system, he progressively develops competencies and masters strategies for rational behavior in the face of uncertainty.

PROGRESS: A preliminary tryout of SIGI was conducted during the summer of 1972. A sample of students used the Introduction and Values system on the PDP-8 and dual-screen terminal and then went on with the Information, Prediction, and Planning systems on the PDP-11 and single-screen terminal. Observations were made for formative evaluation of SIGI. A description and illustration of SIGI appeared in the Summer 1972 issue of the EDUCOM *Bulletin*.

The entire system was then converted to the PDP-11 and single-screen terminal and used by a representative sample of the Mercer County Community College student body during the spring semester of 1973.

A description and illustration of the 1973 version, "Career decision-making: A computer-based system of interactive guidance and information (SIGI)," was published in *Measurement for self-understanding and personal development* (proceedings of 1973 ICTP).

SIGI has since been revised and expanded, and is being reprogrammed to run on multiple terminals.

Validating Effects of SIGI on Career Decision-Making Processes of Students (NSF, 842-4) M. Katz

PURPOSE: To observe the effects of SIGI on the quality and nature of community college students' reasoning and information for career decision-making.

PROGRESS: Criteria were formulated, and interview schedules and other measures were constructed and tried out during the summer of 1972. Revised instruments were used in conjunction with a field test of SIGI at Mercer County Community College during the spring semester of 1973. Evaluative data were collected from experimental and control groups and were published in SIGI: Report of a pilot study under field conditions. A nation-wide field test, evaluation, and dissemination activities are scheduled to continue through 1976.

Departmental Characteristics and Student Academic Growth (844-1)

J. Centra

PURPOSE: (1) To identify departments that have graduates with higher or lower achievement than predicted; and (2) to identify characteristics of these most and least effective departments.

PROGRESS: This study uses the "input-output" model to investigate departmental effects on student cognitive growth in college. Student achievement on the Graduate Record Examinations Advanced Tests or the Undergraduate Program Tests (ETS) were used as the (output) measure of departmental effectiveness; student college admissions scores (SAT or ACT) or their National Merit Scholarship Qualifying Test scores are the input measures. A final report is expected by January 1975.

Development and Validity of an Air Force Vocational Inventory Study (847-1)

G. Echternacht
R. Reilly

PURPOSE: Between January and April 1973, over 3,100 airmen in eight service career fields with at least six months' on-the-job experience and 300 recruits in basic training returned interest inventories, termed the Vocational and Occupational Interest Choice Examination (VOICE). This interest inventory was designed to measure the vocational interests of enlisted men entering the Air Force. Items forming the inventory were primarily generated by examining job analyses in relation to the airman classification structure. The purpose of this effort was to develop and validate a prototype of an interest inventory that could be used by recruiters with the Guaranteed Enlistment Program.

PROGRESS AND MAJOR FINDINGS: Both *a priori* and occupational scales were developed based on responses obtained by mail inventory administration of airmen who indicated satisfaction with their career fields. Scales were developed on half-samples and a cross-validation technique employed. A comparison was made, in order to assess validity, of the number of individuals correctly predicted to be members of a service career field or "men-in-general" using the scales versus the number of individuals one would expect to predict correctly without use of the scales. Scale weights developed in one half-sample were applied to responses obtained in the other half-sample in the above comparison. Recommendations for further developmental effort were made. The reports for this project are found in: (1) *Development and Validity of a Vocational and Occupational Interest Inventory* (AHRL-TR-73-38); (2) *Stage I Development of Voice* (PR-73-43).

Study of Teacher Change in Open Educational Settings E. Chittenden
(Ford Foundation, 858-10)

PURPOSE: This project is a teacher interview study addressed to the problem of identifying and analyzing teachers' perceptions and understandings regarding children, curriculum and the role of the teacher. The interview methodology rests in part on the assumption that a teacher's perception of the classroom and children has pervasive effects on teaching behavior and, in turn, upon the learning environment created for children. Specific goals of the project are as follows: (1) to identify and describe concrete issues and factors which teachers perceive as facilitating and inhibiting of change toward an open classroom; (2) to examine the conceptual characteristics of teachers, and their perceived needs for assistance, in order to gain clearer understanding of the dynamics of change; and (3) in the course of the research, to refine an interview instrument and set of coding procedures which would be applicable to teachers working in innovative reform programs other than open education.

PROGRESS: Sixty teachers in urban public elementary schools have been interviewed. Data from these in-depth interviews are being analyzed.

Student Points of View in Ratings of College Instruction (865-20)

J. Centra
R. Linn

PURPOSE: The purpose of this study was to identify subgroups of students with identifiable "points of view" in their ratings or descriptions of instruction they had received. In addition to identifying dimensions that described the students, the relationship of individual characteristics, such as expected grade in the course and overall grade point average, to those dimensions would be studied.

PROGRESS: This study investigated student points of view in their ratings of specific courses and instructors by analyzing student responses within each of three classes separately, and then sought to generalize the results further by additional analyses with 300 students randomly selected from 402 classes in five colleges.

MAJOR FINDINGS: Different points of view were found for student ratings of course examinations, textbooks and supplementary readings, and class discussions. These various points of view were moderately related to such student characteristics as grades and sex, although not in all three of the classes studied. This last point underscores the importance of the context (i.e., the particular course) in understanding or interpreting the meaning of student ratings.

Test and Job Performance of Various Subgroups within Specified Occupations (Ford Foundation, 879-3)

J. Campbell
L. Crooks
D. Rock
M. Mahoney

PURPOSE: This project was designed to investigate the relationship between job and test performance for different ethnic groups when pertinent background factors are taken into consideration. The project was carried out under a grant from the Ford Foundation as a joint responsibility of ETS and the U. S. Civil Service Commission.

PROGRESS: The relationship of job and test performance for different ethnic groups was studied in three different occupations: medical technician, cartographic technician, and inventory manager. In each of these federal Civil Service occupations, samples of job incumbents were given a battery of aptitude tests selected after careful analysis of job duties and responsibilities. A background, experience, and job task questionnaire was also completed, and superiors rated each incumbent on specially constructed experimental rating scales. Job knowledge tests were also constructed for medical technicians and cartographic technicians and work sample tasks for cartographic technicians and inventory managers. Samples of Black and Caucasian job incumbents were available for all three jobs, and of Mexican-American for the cartographic technicians and inventory manager jobs.

MAJOR FINDINGS: Principal findings include the following:

1. Aptitude tests which have validity in relation to job performance for one ethnic group generally show validity for other ethnic groups as well.

2. Tests which are valid against a rating criterion also show validity against more objective criterion measures.

3. Multiple regression weights determined on a single ethnic group hold up surprisingly well on cross-validation across different ethnic groups.

4. Ethnic group rater-ratee combinations interact to affect the ratings assigned, but the effect appears to be complex and probably differs from one ethnic group to another.

Results of the medical technician project have been reported in Project Reports 69-5, 69-6, 70-16, 70-17, 70-18. Development of the instrumentation has been reported in Project Reports 69-4, 71-22, and 71-23. The proceedings of an invitational conference to discuss and critique the findings has been published, as well as the final report entitled "An investigation of sources of bias in the prediction of job performance--a six year study"--final project report (PR-73-37).

The Effectiveness of Student Feedback in Modifying College Instruction (894-1)

J. Centra

PURPOSE: The effects of formal student feedback to college instructors were investigated in this study through three related questions: Do instructors learn something about their teaching from this feedback? Do instructors improve (i.e., do student ratings improve) after receiving the feedback? And which particular instructional procedures or teacher behaviors tend to be changed?

PROGRESS: Using a specially developed instrument to obtain student reactions to instruction, student feedback was given half way through a semester to one group of instructors (the experimental group) and withheld from a second group (the control group). The instrument was readministered to both groups at the end of the semester as well as to a post hoc group, and comparisons in changes were made. The number of years of teaching experience, sex of the instructor, instructor self-ratings, and subject area of the course were also examined for their relationship to changes.

Data were collected at five diverse colleges during the Fall 1971 semester. Additional ratings were collected at one of the colleges at the end of the Spring 1972 semester.

MAJOR FINDINGS: Changes in instruction occurred after only a half semester for instructors whose self-evaluations were considerably better than their student ratings. Variables such as subject area of the course, sex of the instructor, and number of years the instructor had taught did not distinguish which instructors made changes. However, a wider variety of teachers changed if given more than a half semester of time and if they received some minimal information to help them interpret their ratings. Results are reported in SIR Report #2 (PR-72-16), and in the Journal of Educational Psychology, 1973, 65(3).

Self-Ratings of College Teachers: A Comparison to Student Ratings (894-1) J. Centra

PURPOSE: The purpose of this study was to investigate teachers' self-ratings of instruction by comparing them to ratings given by students. Discrepancies between self-ratings (or self-descriptions) and those provided by students would underscore the need for student feedback as well as highlight specific areas of instruction where feedback is most essential.

PROGRESS: The sample consisted of 343 teaching faculty from five colleges; these teachers, as well as the students in one of their classes, responded to a 21-item instructional report questionnaire. The data were collected during the mid-semester of Fall 1972, and analyses have been completed.

MAJOR FINDINGS: Correlating teacher responses to each item with the mean class responses (across the 343 classes) disclosed a modest relationship between the two sets of evaluation: a median correlation of .21 for the items. In addition to the general lack of agreement between self- and student evaluations, there was also a tendency for teachers as a group to give themselves better ratings than their students did. Comparisons between student and faculty responses were also made across items, and a rank correlation of .77 indicated a good deal of similarity in the way the two groups rank ordered the items.

Discrepancies between individual teacher ratings and ratings given by the class were further analyzed for (a) sex of the teacher (no difference found); (b) number of years of teaching experience (no difference); and (c) subject area of the course (differences noted for natural science courses vs. those in education and applied areas). Among other conclusions, the results of this study would argue for the collection of student ratings to supplement self-ratings. The results are reported in RB-72-33 and as part of SIR Report #2, and in the Journal of Educational Measurement, 1973, 10(4).

Cheating Tests (984-84) F. Lord

PURPOSE: An examinee has taken parallel forms of an aptitude test on three occasions. Two of the scores are so discrepant as to raise a suspicion of cheating. A statistical test is desired for the hypothesis of no cheating.

MAJOR FINDINGS: Under normality assumptions, a uniformly most powerful similar test is derived for this null hypothesis (RM-74-7).